



MASTERARBEIT

Herr
Thomas Hache

**Analyse der Preisentwicklung an
Commoditymärkten mithilfe von
Verfahren der Extremwerttheorie**

2012

MASTERARBEIT

Analyse der Preisentwicklung an Commoditymärkten mithilfe von Verfahren der Extremwerttheorie

Autor:

Thomas Hache

Studiengang:

Diskrete und Computerorientierte Mathematik

Seminargruppe:

ZD10w1-M

Erstprüfer:

Prof. Dr. Egbert Lindner

Zweitprüfer:

Dr. Maximilian Rinck

Mittweida, 2012

I. Inhaltsverzeichnis

Inhaltsverzeichnis	2
1 Einleitung	3
1.1 Die European Commodity Clearing AG	3
1.2 Ziel der Arbeit	4
2 Mathematische Grundlagen	7
2.1 Untersuchung der Daten	7
2.2 Autoregressive Prozesse	10
2.3 Autoregressive Prozesse mit Sprüngen	12
2.4 Stationariät	14
2.5 Value-at-Risk	15
2.6 PP-Plot und QQ-Plot	16
3 Theorie der Blockmaxima	19
3.1 Allgemeine Extremwertverteilungen	19
3.2 Max-Anziehungsbereiche	24
3.3 Satz von Fisher-Tippett	27
3.4 Extremwertverhalten von stationären Zeitreihen	33
3.5 Schätzung der Parameter und des Value-at-Risk	37
3.6 Wahl der Blocklänge	40
3.7 Anwendung des Verfahrens	42
4 Peak over Threshold (POT)	51
4.1 Verallgemeinerte Pareto-Verteilungen	51
4.2 Satz von Pickands-Balkema-de Haan	53
4.3 Wahl des Schwellenwertes	55
4.4 Schätzung der Parameter und des Value at Risk	58
4.5 Anwendung des Verfahrens	62
5 Zusammenfassung und Ausblick	81

A	Kwiatkowski-Phillips-Schmidt-Shin-Test	83
B	Kolmogorov-Smirnov-Test	85
	Literaturverzeichnis	87

1 Einleitung

1.1 Die European Commodity Clearing AG

Die European Energy Exchange (EEX) ist eine der größten Energiebörsen Europas mit Sitz in Leipzig. Sie entstand 2002, indem die deutschen Strombörsen Frankfurt (EEX) und Leipzig (LPX) fusionierten. Als Tochterunternehmen der EEX wurde 2006 die European Commodity Clearing AG (ECC) gegründet, um als Clearinghaus der EEX in börslichen und außerbörslichen Geschäften mit Strom, Gas, Emissionsrechten und Kohle zu dienen. Unter Clearing versteht man das Abwickeln von Geschäften. Dabei geht es um Termin- und Spotmarktgeschäfte, die von der ECC gecleart werden. Das Settlement beschäftigt sich mit der Übereignung von Energiegütern nach dem Geschäftsabschluss. Dabei kann die Übereignung in Form von Energie oder als finanzielle Ausgleichszahlung erfolgen. Diese beiden Prozesse stehen im Vordergrund der Tätigkeiten der ECC. Die European Commodity Clearing AG ist ein zentraler Kontrahent (Central Counterparty). Das heißt, sie tritt als Vertragspartei zwischen Käufer und Verkäufer und benötigt daher eine Banklizenz in Deutschland. In Europa gibt es noch andere Großbanken, die das Clearinggeschäft der Strombörse unterstützen, sogenannte Clearing Mitglieder (CM). Alle Handelsteilnehmer an der Strombörse werden als Non-Clearing Mitglieder (NCM) bezeichnet. Im Verlaufe der letzten 6 Jahre erweiterte die ECC stetig ihr Angebot an Produkten sowie die Anzahl der Teilnehmer und Partnerbörsen. Als zentraler Kontrahent ist sie für das volle Risiko eines möglichen Ausfalls der Handelsteilnehmer verantwortlich. Um dies möglichst zu vermeiden, werden folgende Maßnahmen ergriffen:

1. Die ECC wählt als Clearing Mitglieder (CM), mit denen sie ausschließlich direkte Handelsbeziehungen pflegt, nur Banken aus der EU oder der Schweiz mit einem bestimmten Eigenkapital und einer bestimmten Mindestbonität.
2. Es werden Sicherheitsanforderungen (Margins) für jedes Non-Clearing-Mitglied (NCM) berechnet, welches offene Positionen besitzt. Die Margins gelten als Absicherung gegen ungewöhnliche Preisbewegungen am Markt.
3. Im Falle, dass es zu ungewöhnlichen Preisbewegungen am Markt kommt und ein Handelsteilnehmer dadurch mit seinen Zahlungen in Verzug gerät, existiert der Clearing-Fonds als zusätzliche Absicherung. Alle Clearing Mitglieder sind verpflichtet, in den Clearing-Fonds einzuzahlen.
4. Das bestehende Restrisiko wird mit dem Eigenkapital der ECC abgedeckt.

1.2 Ziel der Arbeit

Die Wahrscheinlichkeitsverteilung von Extremwerten ist in vielen Anwendungsfällen der Finanzwelt von großem Interesse, zum Beispiel bei der Abschätzung von Verlusten oder der Analyse von Preisentwicklungen. Die Abteilung Risk Controlling and Compliance der ECC beschäftigt sich mit der Risikoüberwachung und -steuerung. Dazu gehört neben dem Aufsichtsrecht auch die quantitative Risikosteuerung und Modellierung. Diese beschäftigen sich hauptsächlich mit der Berechnung von Margins. Zur Kalkulation dieser Sicherheitsanforderungen wird eine in der Finanzwelt weit verbreitete Größe herangezogen: der Value-at-Risk, kurz *VaR*. Dieser Wert gibt den Verlust an, der mit einer Wahrscheinlichkeit von $1 - \alpha$, $0 < \alpha < 1$ in einer bestimmten Periode t nicht überschritten wird. Mathematisch gesehen ist der $VaR_{1-\alpha}^t$ das $1 - \alpha$ -Quantil der zugrunde liegenden Verteilung, was in Kapitel 2.5 noch genauer erläutert wird. Die Schätzung des Value-at-Risk ist jedoch in der Praxis oft sehr schwierig, da die Verteilung der Daten, insbesondere bei multivariaten Problemen unbekannt ist. Bei der ECC wird der Value-at-Risk berechnet, indem man die relativen Preisänderungen als normalverteilt annimmt. Das hat jedoch den Nachteil, dass das Auftreten von Extremwerten unterschätzt wird. Aus diesem Grund wird bei der ECC ein pauschaler Sicherheitsaufschlag von 50% auf den *VaR* veranschlagt, welcher jedoch dazu führt, dass in den meisten Fällen das Risiko deutlich überschätzt wird und die Margins deshalb höher kalkuliert werden, als es notwendig wäre. Diese hohen Sicherheitsanforderungen sind für Handelsteilnehmer ein Problem, da diese übermäßige Geldbeträge bereitstellen müssen, welche sie unter Umständen nicht besitzen.

Ziel der vorliegenden Arbeit ist, die dokumentierten extremen Ereignisse mithilfe von extremwerttheoretischen Verfahren zu analysieren, um mit den gewonnenen Ergebnissen den *VaR* besser abschätzen zu können. Dabei werden zwei Methoden vorgestellt und an Beispielen und realen Daten angewendet. Als erstes beschäftigen wir uns mit der Theorie der Blockmaxima, bei der die Datenreihe in zeitgleiche Blöcke eingeteilt und anschließend das Maximum aus jedem Block gezogen wird. Zunächst erfolgt ein Überblick über allgemeine Extremwertverteilungen. Es wird erklärt, welche Verteilungen gegen Extremwertverteilungen konvergieren und der Satz von Fisher-Tippett, der Stützpfeiler der Blockmaxima-Theorie, wird vorgestellt und bewiesen. Ein wichtiger Vorgang der Verteilungsanalyse ist die Schätzung der Parameter, weshalb eine ausführliche Beschreibung der Maximum-Likelihood-Schätzung in diesem Kapitel erfolgt. Interessant ist ebenfalls die Wahl der Blocklänge, weswegen dieses Problem im Verlauf des Abschnitts analysiert wird. Anschließend folgen mehrere Anwendungsbeispiele mit dem Programm MATLAB.

Als zweites Verfahren stellen wir die Theorie der Schwellenwertüberschreitungen vor (Peak over Threshold). Betrachtet werden hier nur die Überschreitungen einer Schwelle, denen allgemeine Pareto-Verteilungen zugrunde liegen. Als zentraler Gegenstand der Peak over Threshold dient der Satz von Pickands-Balkema-de Haan. Ein weiterer wichtiger Punkt ist die Wahl des Schwellenwertes, über den in diesem Kapitel gespro-

chen wird. Wie bei der Theorie der Blockmaxima müssen auch bei Peak over Threshold die Parameter der allgemeinen Pareto-Verteilung geschätzt werden. Hierbei bieten sich gleich mehrere Schätzmethoden an, die ausführlich diskutiert werden. Anschließend folgen Anwendungsbeispiele mit MATLAB, die die Arbeitsweise der Methode verdeutlichen. Alle gewonnenen Erkenntnisse werden im letzten Kapitel zusammengefasst. Als Abschluss erfolgt ein Ausblick über weitere mögliche Betrachtungen.

2 Mathematische Grundlagen

Bevor wir mit unseren Ausführungen starten können, müssen noch einige Aspekte vorweg genommen werden. In diesem Kapitel beschäftigen wir uns unter anderem mit den Zeitreihendaten, auf die wir später extremwerttheoretische Verfahren anwenden wollen. Uns interessiert wie sich die Zeitreihen Verhalten, wie Ihre Autokorrelation aussieht und ob Sie stationär sind oder nicht. Des Weiteren wollen wir versuchen, die gegebenen Zeitreihen mit bestimmten stochastischen Prozessen nachzuahmen, um später einen Überblick zu bekommen, wie sich die vorgestellte Extremwerttheorie auf theoretischen Daten verhält, bei denen alle Voraussetzungen der Extremwerttheorie (siehe Kapitel 3.4) erfüllt sind. Darüber hinaus führen wir am Ende dieses Kapitels den Value-at-Risk als ein populäres Risikomaß ein und diskutieren wichtige Aspekte seiner Berechnung.

2.1 Untersuchung der Daten

Die Zeitreihendaten, mit denen wir uns in dieser Arbeit beschäftigen werden, stammen vom sogenannten Phelix-Index. Dies ist die Abkürzung für Physical Electricity Index und ist der Börsenindex für Strom aus Deutschland und Österreich. Er stellt den Durchschnittspreis der Ware Strom am Spotmarkt an einem Tag dar und ist maßgeblicher Basiswert für die finanziellen Termingeschäfte. Man unterscheidet Phelix-Baseload (Grundlast) und Phelix-Peakload (Spitzenlast), wobei sich die beiden Begriffe auf unterschiedliche Tageszeiten beziehen. Während der Baseload den Durchschnittspreis aller Stundenkontrakte eines kompletten Tages abbildet, gehen beim Peakload nur die Stunden von 08:00 bis 20:00 Uhr in den Preis ein, also die Hochauslastung des Stromnetzes.

Die Zeitreihen des Phelix-Index, genannt F1B (Base) und F1P (Peak), haben in dieser Arbeit eine Länge von 2662 Börsentagen und beziehen sich auf den Zeitraum vom 08.02.2005 bis 25.05.2012:

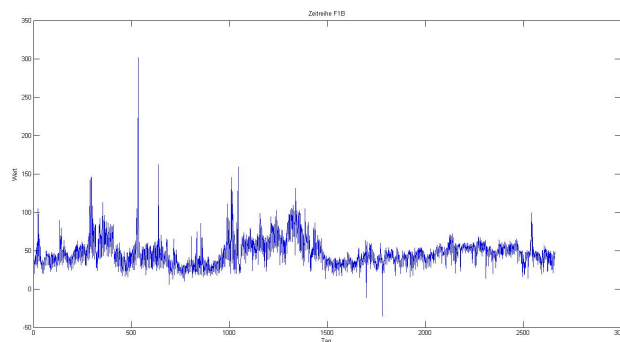


Abbildung 2.1: Phelix-Index im Base mit 2662 Börsentagen

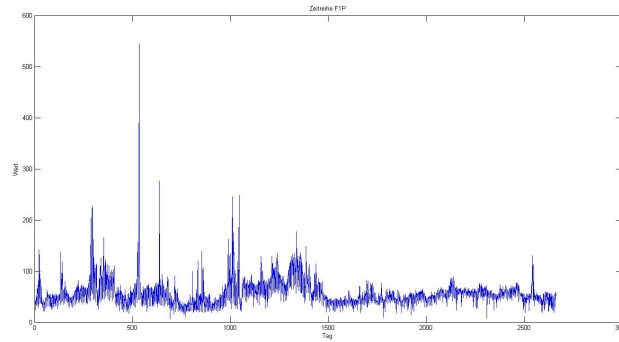


Abbildung 2.2: Phelix-Index im Peak mit 2662 Börsentagen

Auffällig in den beiden Abbildungen 2.1 und 2.2 ist der Spitzenwert nach mehr als 500 Börsentagen. Es handelt sich dabei um den 27.07.2006, der heißeste Tag im diesem Jahr. Der Grund für den extremen Strompreis war, dass durch die Hitzewelle wassergekühlte Großkraftwerke ihre Stromproduktion drosseln mussten und damit für massiv steigende Preise sorgten. Das andere Extremum kann nach ca. 1700 Börsentagen am 05.10.2009 beobachtet werden. Dort kam es an der Leipziger EEX in der Nacht erstmals zu negativen Strompreisen. Als Grund dafür gaben Stromhändler die schwache Nachfrage um diese Zeit und ein üppiges Windangebot bei nicht sonderlich tiefen Temperaturen an. Somit war es für ein paar Stunden möglich, Strom einzukaufen und dabei noch Gewinn zu machen. Flexible Kraftwerke, die ihre Anlagen bei Bedarf herunterfahren können, profitierten von diesem seltenen Phänomen. Dagegen mussten unflexible Kernkraftwerke bei diesem Geschäft draufzahlen.

Wir haben gesehen, dass der Strompreis mitunter kurzen, aber dafür extremen Schwankungen unterliegt. Gründe dafür können das Wetter, ein Kraftwerksausfall oder auch politische Bestimmungen sein. Interessant ist nun, welche Abhängigkeitsstruktur in den Zeitreihen steckt.

Definition 2.1 Sei X_t ein stochastischer Prozess. Dann bezeichnen wir

$\mu(t) := E(X_t)$ als **Mittelwertfunktion**,

$\sigma^2(t) := D^2(X_t)$ als **Varianzfunktion** und

$\gamma(\tau) := Cov(X_{t+\tau}, X_t)$ als **Kovarianzfunktion**.

Definition 2.2 Sei Y_t eine Zeitreihe und $\tau \geq 0$. Die **Autokorrelation** von Y_t ist definiert durch [10]

$$Corr(Y_{t+\tau}, Y_t) = \frac{Cov(Y_{t+\tau}, Y_t)}{\sqrt{D^2(X_{t+\tau})D^2(X_t)}}. \quad (2.1)$$

Definition 2.3 Die **empirische Kovarianzfunktion** c_τ einer Zeitreihe Y_t , sowie ihre **em-**

pirische Autokorrelationsfunktion, kurz ACF, r_τ sind definiert durch

$$c(\tau) = \frac{1}{N} \sum_{t=1}^{N-\tau} (y_t - \bar{y})(y_{t+\tau} - \bar{y}), \quad \tau > 0 \quad (2.2)$$

$$r(\tau) = \frac{c_\tau}{c_0}. \quad (2.3)$$

Mit τ werden die sogenannten Lags bezeichnet [10].

Wenden wir nun Gleichung (2.2) und (2.3) auf unsere Zeitreihendaten an, ergeben sich die folgenden Korrelogramme:

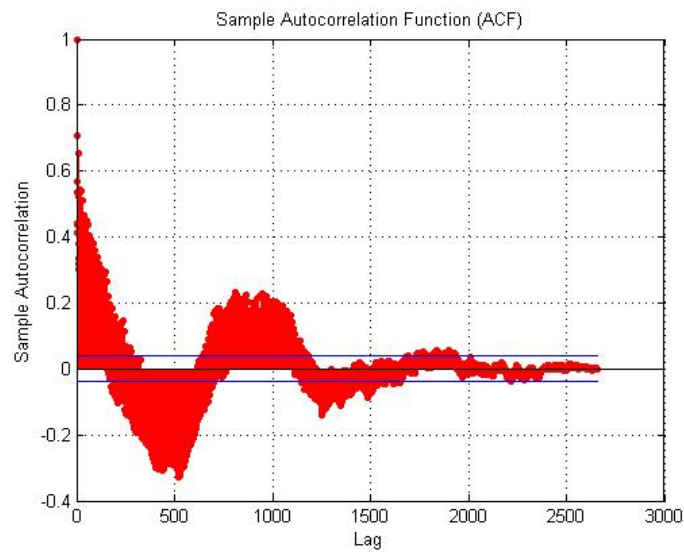


Abbildung 2.3: Autokorrelation des Baseload über 2662 Börsentage

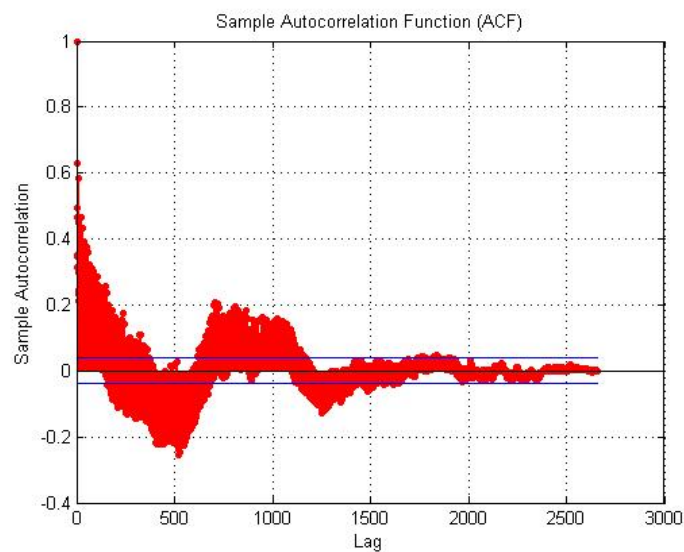


Abbildung 2.4: Autokorrelation des Peakload über 2662 Börsentage

Man erkennt sowohl in Abbildung 2.3, als auch in Abbildung 2.4 deutlich, dass die Abhängigkeit eines Datenpunkts zum Anfangspunkt mit der Zeit abnimmt.

Als ein weiterer interessanter Aspekt erweist sich die Frage, ob wir diese Zeitreihen mit stochastischen Modellen nachahmen können, da unser Datenumfang mit 2662 Datenpunkten eher klein ist. Falls dem so wäre, hätten wir die Möglichkeit beliebig viele Zeitreihenschritte zu simulieren. Da unsere Zeitreihendaten eine schwache Abhängigkeitsstruktur aufweisen, werden wir versuchen den F1P und F1B mit autoregressiven Prozessen nachzuahmen.

2.2 Autoregressive Prozesse

Ein wichtiger Bestandteil autoregressiver Prozesse ist das weiße Rauschen oder auch White-Noise-Prozess genannt.

Definition 2.4 Ein **White-Noise-Prozess** oder **weißes Rauschen** ist eine Folge von unabhängigen identisch verteilten Zufallsgrößen. Im Folgenden werden wir solche Prozesse mit ε_t bezeichnen [10].

In dieser Arbeit beschränken wir uns auf das spezielle weiße Rauschen mit Erwartungsfunktion $E(U_t) = 0$ und Varianzfunktion $D^2(U_t) = \sigma^2$. Damit sind wir nun in der Lage autoregressive Prozesse zu definieren.

Definition 2.5 Wir nennen einen stochastischen Prozess X_t einen **autoregressiven Prozess der Ordnung** p , kurz $AR(p)$, wenn er für alle Konstanten $c, \phi_1, \dots, \phi_p \in \mathbb{R}$ und weißes Rauschen ε_t („Innovationen“) mit Varianzfunktion σ^2 folgender Gleichung genügt [10]:

$$X_t = c + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t, \quad t \in \mathbb{Z}.$$

Uns interessieren hier nur die autoregressiven Prozesse 1. Ordnung, also $AR(1)$ mit $c = 0$, $\sigma^2 = 1$ und $\phi_1 = \alpha$, wobei $0 < \alpha < 1$:

$$X_t = \alpha X_{t-1} + \varepsilon_t, \quad t \in \mathbb{Z}. \quad (2.4)$$

Die Simulation eines solchen Prozesses ist recht einfach, da man nur die Bildungsvorschrift (2.4) befolgen muss:

Algorithmus 1 Simulation eines $AR(1)$ -Prozess

```

Setze  $X_1$  auf eine  $N(0, 1)$ -verteilte Zufallszahl
for  $i = 2$  to  $T$  do
  Setze  $\varepsilon_i$  auf eine  $N(0, 1)$ -verteilte Zufallszahl
  Setze  $X_i = \alpha X_{i-1} + \varepsilon_i$ 
end for

```

Nach Anwendung von Algorithmus 1 ergab sich die folgende Zeitreihe:

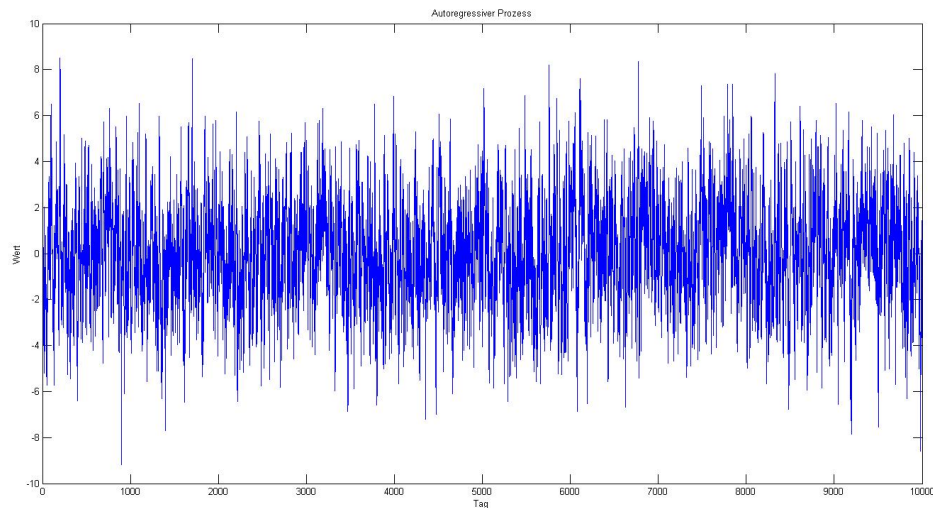


Abbildung 2.5: Autoregressiver Prozess mit 10.000 Schritten

Bei autoregressiven Prozessen ist allgemein bekannt, dass Sie eine schwache Abhängigkeitsstruktur aufweisen. Man kann zeigen, dass für die Kovarianzfunktion eines $AR(1)$ -Prozesses folgendes gilt [10]:

$$\gamma(\tau) = \text{Cov}(X_{t+\tau}, X_t) = \alpha^{|\tau|} \gamma(0).$$

Somit erhält man für die Autokorrelationen eines $AR(1)$ -Prozesses [10]:

$$\rho(\tau) = \frac{\gamma(\tau)}{\gamma(0)} = \alpha^{|\tau|}.$$

Für $\tau \rightarrow \infty$ konvergieren also die Autokorrelationen $\rho(\tau)$ gegen 0:

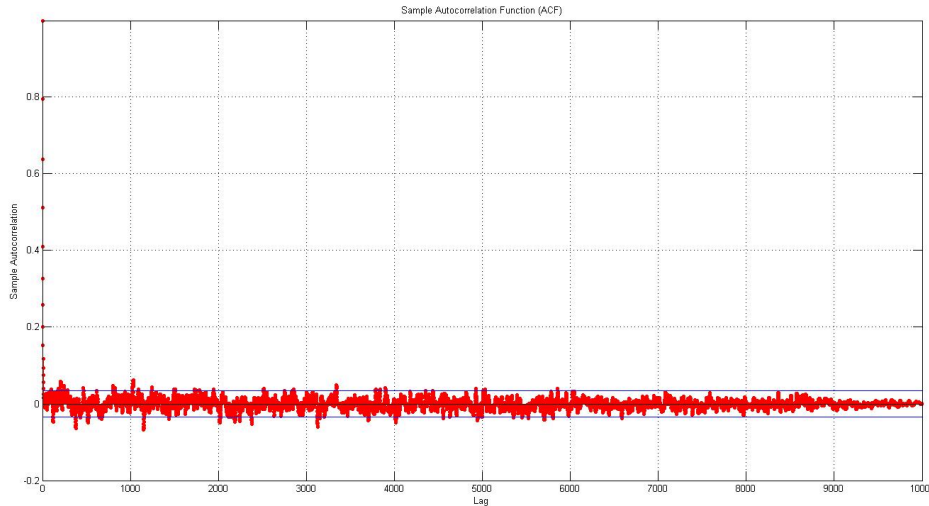


Abbildung 2.6: Autokorrelation des autoregressiven Prozesses über 10.000 Schritte

Bei näherer Betrachtung der simulierten Zeitreihe stellen wir fest, dass in diesem Verlauf keine extremen Ereignisse auftreten, die sich deutlich von den anderen abheben. Auch in der Energiebranche ist es üblich, wie wir im vorherigen Abschnitt sehen konnten, dass hin und wieder extreme Preisspitzen auftreten. Da wir uns genau für solche extremen Realisierungen interessieren, müssen noch weitere Überlegungen getätigt werden.

2.3 Autoregressive Prozesse mit Sprüngen

Um seltene extreme Preisspitzen simulieren zu können, greifen wir auf die Theorie des zusammengesetzten Poisson Prozesses zurück. Ein zusammengesetzter Poisson Prozess ist ein stochastischer Prozess, bei dem die Wartezeit zwischen zwei Sprüngen exponentialverteilt, aber die Sprunghöhe beliebig verteilt ist.

Definition 2.6 Sei N_t ein Poisson Prozess mit Parameter $\lambda > 0$ und $(Y_i)_{i \geq 1}$ eine Folge von unabhängigen, identisch verteilten Zufallsgrößen mit Verteilung F . Der Prozess

$$X_t = \sum_{i=1}^{N_t} Y_i \quad (2.5)$$

heißt **zusammengesetzter Poisson Prozess** ([13] und [12]).

Die Idee ist nun die Modelle (2.4) und (2.5) miteinander zu vereinen. Um das zu gewährleisten, addieren wir einfach beide Prozesse und erhalten einen neuen autoregressiven Prozess mit gekoppelten Poisson Prozess:

$$X_t = \alpha X_{t-1} + \varepsilon_t + \sum_{i=1}^{N_t} Y_i \quad (2.6)$$

Wir wählen die Zufallsgrößen Y_i als standard normalverteilt, da es, wie wir gesehen haben, positive und negative Preisspitzen geben kann.

Simulationsvorschrift:

Algorithmus 2 Simulation eines $AR(1)$ -Prozess mit gekoppelten Poisson Prozess

Setze N_t auf eine $Pois(\lambda T)$ -verteilte Zufallszahl

Simuliere N_t gleichverteilte Zufallszahlen $(U_j)_{j=1}^{N_t}$ auf $[0, T]$

Simuliere N_t unabhängige $N(0, 1)$ -erteilte Zufallszahlen $(Y_j)_{j=1}^{N_t}$

Setze X_1 auf eine $N(0, 1)$ -verteilte Zufallszahl

for $i = 2$ to T **do**

 Setze ε_i auf eine $N(0, 1)$ -verteilte Zufallszahl

 Setze $X_i = \alpha X_{i-1} + \varepsilon_i + \sum_{j=1}^{N_t} Y_j 1_{U_j \leq t}$

end for

Nach Anwendung von Algorithmus 2 ergab sich der folgende Verlauf:

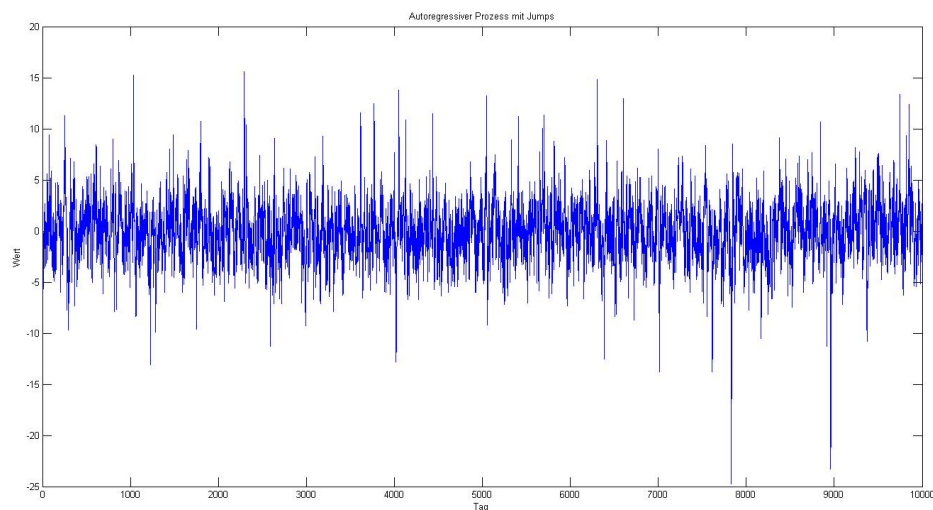


Abbildung 2.7: Autoregressiver Prozess mit Sprüngen und 10.000 Schritte

An diesem Modell ist ersichtlich, dass den extremen Ereignissen mehr Aufmerksamkeit gewidmet wird, als mit einfachen $AR(1)$ -Prozessen. Aber an der Abhängigkeitsstruktur hat sich nichts verändert:

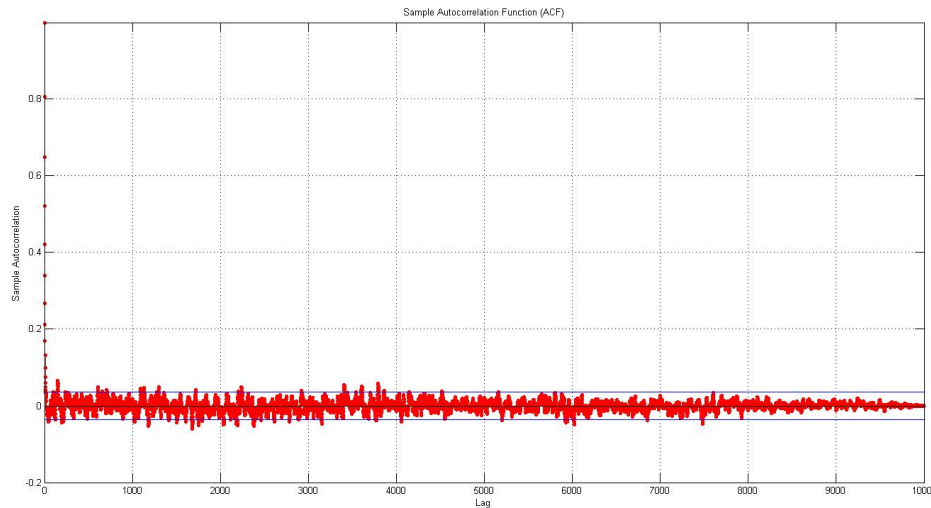


Abbildung 2.8: Autokorrelation des autoregressiven Prozesses mit Sprüngen über 10.000 Schritte

2.4 Stationariät

Zunächst müssen wir definieren, was im weiteren Verlauf der Arbeit unter Stationarität verstanden werden soll [10].

Definition 2.7 Ein stochastischer Prozess X_t heißt

- **mittelwertstationär**, falls $\mu(t) = \mu \quad \forall t$
- **varianzstationär**, falls $\sigma^2(t) = \sigma^2 \quad \forall t$
- **kovarianzstationär**, falls $Cov(X_{t+\tau}, X_t) = Cov(X_\tau, X_1) \quad \forall t$.

X_t heißt **schwach stationär**, falls X_t mittelwert- und kovarianzstationär ist.

Für unsere späteren Untersuchungen reicht es, wenn die Zeitreihen die Eigenschaften der schwachen Stationarität erfüllen. Wenn wir in späteren Kapiteln also von Stationarität sprechen, so ist immer schwache Stationarität gemeint.

Nun interessiert uns, ob die Preiszeitreihen F1B und F1P, sowie die vorhin eingeführten autoregressiven Prozesse stationär sind.

Bemerkung 2.8 Man kann zeigen, dass ein $AR(1)$ -Prozess genau dann stationär ist, wenn $|\alpha| < 1$.

Beweis: siehe [10]. □

Wir unterziehen nun die Zeitreihen dem Kwiatkowski-Phillips-Schmidt-Shin-Test, kurz KPSS-Test, einem Hypothesentest, der direkt auf Stationarität testet (siehe Anhang A).

Nach Anwendung dieses Tests erhalten wir folgende Kenngrößen:

Zeitreihentyp	KPSS-Statistik	99%-Quantil der KPSS-Verteilung
F1B	0,4565	0,76174
F1P	0,8326	0,76174
$AR(1)$ -Prozess	0,6891	0,76174
$AR(1)$ -Prozess mit Jumps	0,1885	0,76174

Wenn die Teststatistik das 99%-Quantil überschreitet, so muss die Stationaritätsannahme abgelehnt werden. Dies ist jedoch nur bei der Preiszeitreihe F1P der Fall. Dass F1P nicht stationär ist, ist jedoch kein „K.O.“-Kriterium. Da Strompreise saisonalen Trends unterliegen sind die Zeitreihen innerhalb kurzer Perioden im Allgemeinen nicht stationär. Da der Spotpreis aber stets auf ein mittleres Niveau zurückkehrt [12] und in der Extremwerttheorie die globalen Ereignisse einer Zeitreihe betrachtet werden, können wir annehmen, dass der vollständige Verlauf einer Zeireihe wieder stationär ist. Wir werden in einem späteren Kapitel feststellen, dass die Stationarität nur eine hinreichende Bedingung ist. Hierzu werden im Kapitel 3.4 noch nähere Erklärungen folgen.

2.5 Value-at-Risk

Eines der bedeutendsten Risikomaße in der Finanzwirtschaft ist der Value-at-Risk. Dieser gibt an, wie hoch der Verlust innerhalb einer bestimmten Zeitspanne ist, der mit einer gegebenen Wahrscheinlichkeit nicht überschritten wird. Im Folgenden wollen wir diesen Sachverhalt herleiten. Zunächst betrachten wir die einfache Wahrscheinlichkeitsverteilung von Realisierungen einer Zufallsgröße X , die den Verlust beschreibt:

$$P(X \leq x) =: F(x).$$

Gesucht ist nun der Wert x_α , der mit einer Wahrscheinlichkeit von $1 - \alpha = P(X \leq x)$, mit $0 < \alpha < 1$, nicht überschritten wird:

$$1 - \alpha = F(x_\alpha).$$

Nun stellt man nach x_α um und erhält:

$$F^{-1}(1 - \alpha) = x_\alpha = VaR_{1-\alpha}.$$

Der Value-at-Risk ist also das $1 - \alpha$ -Quantil der zugrunde liegenden Verteilung F .

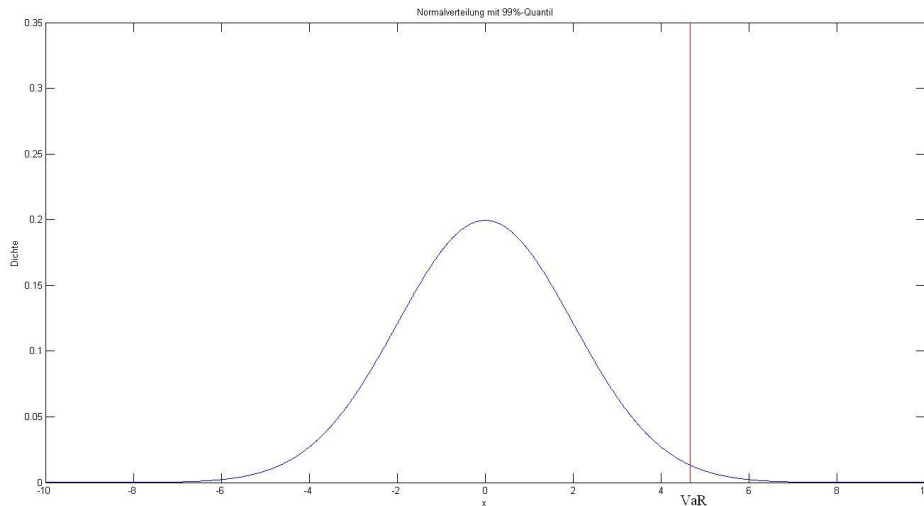


Abbildung 2.9: Der $VaR_{99\%}$ am Beispiel der Normalverteilung $N(0, 2)$

Falls die Quantilfunktion einer Verteilung nicht existiert, so muss auf die verallgemeinerte Inverse einer Verteilungsfunktion zurückgegriffen werden [1]:

$$F^{-1}(1 - \alpha) = \inf\{x : F(x) \geq 1 - \alpha\}, \quad \forall x \in \mathbb{R}. \quad (2.7)$$

Da F aber im allgemeinen unbekannt ist muss zur Berechnung des $VaR_{1-\alpha}$ auf empirische Schätzmethoden zurückgegriffen werden, auf die wir in Kapitel 4.3 zu sprechen kommen.

2.6 PP-Plot und QQ-Plot

Wir werden im Verlauf der Arbeit auf das Problem der Anpassungsgüte stoßen. Nachdem eine Verteilung \hat{F} an die Daten angepasst ist, muss untersucht werden, wie gut oder wie schlecht diese Anpassung ist. Dafür werden nun zwei Ideen vorgestellt:

Zunächst sortieren wir die Daten X_1, \dots, X_n der Größe nach und erhalten die Ordnungsstatistik:

$$X_{(1)} \leq \dots \leq X_{(n)}.$$

Mit der Ordnungsstatistik kann man nun den PP-Plot definieren [5]:

Definition 2.9 Der **PP-Plot** ist die Menge

$$\left\{ \left(\hat{F}(X_{(i)}), \frac{i}{n+1} \right) : i = 1, \dots, n \right\} \subset [0, 1]^2.$$

Trifft die Aussage zu, dass X_i gemäß \hat{F} verteilt ist, so sollte $\hat{F}(X_i) \approx \frac{i}{n+1}$ sein bzw. die Punkte eines PP-Plots auf der Winkelhalbierenden liegen, wie folgende Abbildung zeigt:

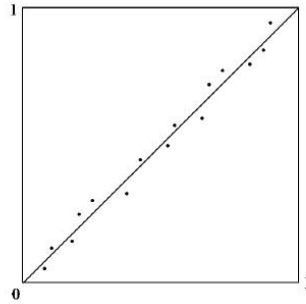


Abbildung 2.10: Beispiel eines PP-Plot (Probability Plot) [1]

Bei dieser Methode gibt es jedoch einen Nachteil. Für $i = n$ gilt $\hat{F}(X_i) \approx 1$ und $\frac{i}{n+1} \approx 1$ unabhängig davon, ob \hat{F} präzise an die Daten angepasst wurde oder nicht (siehe Abbildung 2.10). Das heißt, auch wenn \hat{F} im Bereich der großen Werte stark von Daten abweicht, wird das im PP-Plot nicht erkannt. Wir interessieren uns jedoch gerade für die Verteilung der Daten im höheren Wertebereich, sodass auf eine andere Methode zurückgegriffen wird [5]:

Definition 2.10 Der **QQ-Plot** ist die Menge

$$\left\{ \left(\hat{F}^{-1}\left(\frac{i}{n+1}\right), X_{(i)} \right) : i = 1, \dots, n \right\} \in \mathbb{R}^2.$$

Hier werden die $\frac{1}{n+1}, \frac{2}{n+1}, \dots, \frac{n}{n+1}$ -Quantile der Verteilung \hat{F} auf der x -Achse abgebildet und die Ordnungsstatistiken $X_{(1)} \leq \dots \leq X_{(n)}$ auf der y -Achse. Wenn \hat{F} die Daten gut beschreibt, so sollte $\hat{F}^{-1}\left(\frac{i}{n+1}\right) \approx X_{(i)}$ gelten, bzw. sollten die Punkte im QQ-Plot auf der Winkelhalbierenden liegen.

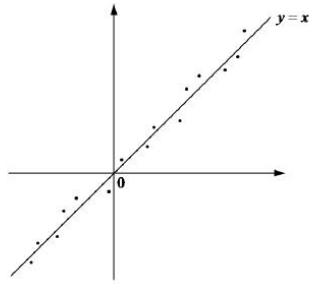


Abbildung 2.11: Beispiel eines QQ-Plot (Quantil-Plot) [1]

In der späteren Anwendung der Verfahren der Extremwerttheorie werden wir immer den QQ-Plot verwenden, um zu sehen, ob die Quantile der geschätzten Verteilung mit den empirischen Quantilen harmonieren.

3 Theorie der Blockmaxima

Als erstes extremwerttheoretisches Verfahren wollen wir uns mit der Theorie der Blockmaxima beschäftigen. Die Idee hinter dieser Methode ist das Aufteilen der einzelnen Beobachtungen in mehrere gleichgroße, disjunkte Blöcke, sodass die Vereinigung aller Blöcke wieder die Menge der Ausgangsdaten ergibt. Anschließend wird aus jedem Block das Maximum gezogen. Dabei erhält man eine neue Datengrundlage, an die, getrennt von den Ausgangsdaten, eine Extremwertverteilung angepasst wird. Das Aussehen und die Struktur einer solchen Extremwertverteilung werden wir im folgenden Kapitel beschreiben.

3.1 Allgemeine Extremwertverteilungen

Wir betrachten zunächst folgende Definitionen [1]:

Definition 3.1 Seien X_1, X_2, \dots unabhängige, identisch verteilte Zufallsgrößen. Unter dem **Maximum** M_n verstehen wir den folgenden Ausdruck:

$$M_n := \max\{X_1, \dots, X_n\}, n \in \mathbb{N}. \quad (3.1)$$

Definition 3.2 Die **Verteilungsfunktion** F und die **Tailfunktion** \bar{F} von X_i sind definiert durch

$$F(t) := P(X_i \leq t), \bar{F}(t) = 1 - F(t) = P(X_i > t), t \in \mathbb{R}. \quad (3.2)$$

Auf Grundlage dieser zwei Definitionen ergibt sich [14]:

Satz 3.3 Die Verteilungsfunktion von M_n ergibt sich aus

$$P(M_n \leq t) = F^n(t), \forall t \in \mathbb{R}. \quad (3.3)$$

Beweis: Aus Gleichung (3.1) und (3.2) folgt

$$P(M_n \leq t) = P(\max\{X_1, \dots, X_n\} \leq t) = P(X_1 \leq t, \dots, X_n \leq t).$$

Da laut Voraussetzung von Gleichung (3.1) die einzelnen Zufallsgrößen X_i unabhängig sind, ergibt sich schließlich

$$P(X_1 \leq t, \dots, X_n \leq t) = P(X_1 \leq t) \cdot \dots \cdot P(X_n \leq t) = F^n(t).$$

□

Definition 3.4 Der rechte Endpunkt einer Verteilungsfunktion F ist definiert durch [14]

$$x^* := \sup\{t \in \mathbb{R} | F(t) < 1\}. \quad (3.4)$$

Dabei kann der Fall eintreten, dass $x^* = \infty$.

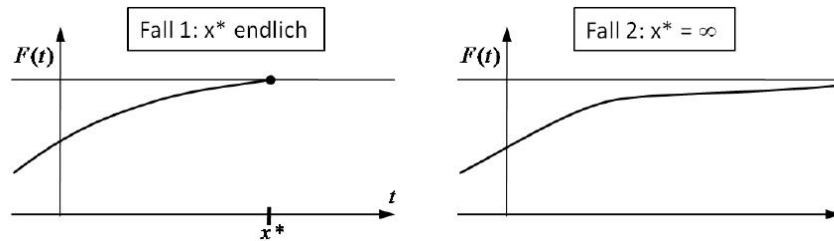


Abbildung 3.1: Veranschaulichung des rechten Endpunkts x^* einer Verteilungsfunktion F

Die Frage, die uns somit interessiert, lautet: Wie verhält sich M_n für $n \rightarrow \infty$?

Satz 3.5 Für $n \rightarrow \infty$ konvergiert die Zufallsgröße M_n fast sicher ($\xrightarrow{f.s.}$) gegen den rechten Endpunkt x^* der Verteilungsfunktion F .

Beweis: Nachzulesen in [14]. □

Somit haben wir eine erste Vorstellung unserer gesuchten Verteilung für Maxima. Dieses Ergebnis gibt uns jedoch kaum eine präzise Vorstellung darüber, wie Extremwertverteilungen aussehen (siehe Abbildung 3.2), da die Verteilungsfunktion F der Daten im Allgemeinen unbekannt ist.

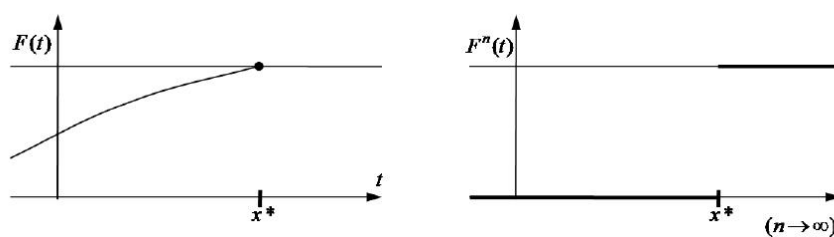


Abbildung 3.2: Veranschaulichung der Verteilungsfunktion von M_n

Die Funktion in Abbildung 3.2 ist für $n \rightarrow \infty$ eine sogenannte „ausgeartete“ Wahrscheinlichkeitsverteilung, für die wir uns nicht interessieren, da diese keine stetige Verteilungsfunktion darstellt.

Das Resultat von Satz 3.5 gibt uns also noch keine ausreichende Vorstellung davon, wie sich die Verteilungsfunktion von M_n für $n \rightarrow \infty$ verhält. Ein ähnliches Problem besteht, wenn man das schwache Gesetz der Großen Zahlen betrachtet. Dieses besagt, dass für

unabhängige, identisch verteilte Zufallsgrößen X_1, X_2, \dots mit endlichem Erwartungswert folgende Wahrscheinlichkeitskonvergenz (\xrightarrow{p}) gilt:

$$\frac{\sum_{i=1}^n X_i}{n} \xrightarrow{p} E(X_1), \quad n \rightarrow \infty.$$

Hierbei erhalten wir ebenfalls keine präzise Vorstellung der Verteilungsfunktion von $\sum_{i=1}^n X_i$. Erst durch den zentralen Grenzwertsatz können genaue Aussagen über die Verteilungsfunktion von $\sum_{i=1}^n X_i$ getroffen werden. Dieser besagt, dass für unabhängige, identisch verteilte Zufallsgrößen X_1, X_2, \dots mit endlichem Erwartungswert μ und endlicher Varianz σ^2 die Verteilungskonvergenz (\xrightarrow{d}) gilt:

$$\frac{\sum_{i=1}^n X_i - \mu}{\sigma\sqrt{n}} \xrightarrow{d} N(0, 1), \quad n \rightarrow \infty.$$

Nun stellt sich folgende Frage: Existiert eine ähnliche Aussage bezüglich $\max\{X_1, \dots, X_n\}$? Um die Antwort zu finden, präzisieren wir unsere Frage:

Seien X_1, X_2, \dots unabhängige, identisch verteilte Zufallsgrößen mit Verteilungsfunktion F . Existieren Folgen von Konstanten $(a_n) \neq 0$ und $(b_n) \in \mathbb{R}$ und eine Verteilungsfunktion G , so dass für $n \rightarrow \infty$

$$\frac{\max\{X_1, \dots, X_n\} - b_n}{a_n} \xrightarrow{d} G \tag{3.5}$$

gilt?

Mit Hilfe der Verteilungskonvergenz mit den Normierungskonstanten a_n und b_n kann somit eine Verteilungsfunktion G beschrieben werden,

$$\lim_{n \rightarrow \infty} P\left(\frac{\max\{X_1, \dots, X_n\} - b_n}{a_n} \leq t\right) =: G(t),$$

für alle $t \in \mathbb{R}$, an denen G stetig ist.

Definition 3.6 Eine Verteilungsfunktion G , für die (3.5) bei irgendeiner Wahl der Normierungskonstanten a_n und b_n , sowie der Verteilungsfunktion F gilt, heißt **Extremwertverteilung**. Umgekehrt gilt, falls (3.5) gilt, so sagen wir, dass F im **Max-Anziehungsbereich** von G liegt [14].

Bemerkung 3.7 Man kann durch ungünstige Wahl von a_n und b_n immer erreichen, dass

(3.5) erfüllt ist. Dabei stoßen wir aber auf eine „ausgeartete“ Verteilungsfunktion G der Form

$$G(t) = \begin{cases} 0, & t < t_0 \\ 1, & t \geq t_0 \end{cases}, \quad (3.6)$$

wobei die Zufallsgröße, die gemäß G verteilt ist, nur den Wert t_0 mit Wahrscheinlichkeit 1 annimmt. Dies erhält man zum Beispiel, indem man $b_n = 0$ setzt und a_n so groß wählt, dass M_n/a_n gegen 0 in Verteilung konvergiert. Aus diesem Grund schließen wir im weiteren Verlauf der Arbeit diese triviale Form von Verteilungsfunktion immer aus.

Mit der obigen Definition haben wir jedoch noch nicht die Existenz von nichttrivialen Extremwertverteilungen bewiesen. Momentan können weder über das Aussehen, noch über die Struktur einer solchen Verteilungsfunktion Informationen gewonnen werden. Im Folgenden werden wir drei verschiedene Extremwertverteilungen herleiten.

Satz 3.8 Seien X_i exponentialverteilt mit Parameter 1 $\forall i$, d.h. $F(t) = 1 - \exp -t$, $t \geq 0$. Dann gilt für die Blockmaxima M_n

$$\lim_{n \rightarrow \infty} P(M_n - \log n \leq t) = e^{-e^{-t}} =: G(t), \quad t \in \mathbb{R}.$$

Das heißt, die Zufallsgröße $M_n - \log n$ konvergiert in Verteilung gegen G [1].

Beweis: Sei $t \in \mathbb{R}$. Mit Gleichung (3.3) gilt

$$P(M_n - \log n \leq t) = P(M_n \leq t + \log n) = F^n(t + \log n).$$

Da die Zufallsgrößen X_i exponentialverteilt sind mit Parameter 1 $\forall i$ und außerdem $t + \log n > 0$, wenn n groß genug, ergibt sich somit:

$$P(M_n - \log n \leq t) = \left(1 - e^{-(t + \log n)}\right)^n = \left(1 + \frac{-e^{-t}}{n}\right)^n \rightarrow e^{-e^{-t}}, \quad n \rightarrow \infty.$$

Damit erhalten wir $M_n - \log n \xrightarrow{P} G$, für $n \rightarrow \infty$. □

Definition 3.9 Die Verteilungsfunktion $G(t) = e^{-e^{-t}}$ heißt **Gumbel-Verteilung** [1].

Erinnern wir uns an (3.5), dann wissen wir, dass die Exponentialverteilung gemäß Satz 3.8 im Max-Anziehungsbereich der Gumbel-Verteilung liegt.

Satz 3.10 Seien X_i Pareto-verteilt mit Parameter $\alpha > 0$, das heißt:

$$F(t) = \begin{cases} 1 - \frac{1}{t^\alpha}, & t \geq 1 \\ 0, & t \leq 1. \end{cases}$$

Dann gilt

$$\lim_{n \rightarrow \infty} P\left(\frac{M_n}{n^{\frac{1}{\alpha}}} \leq t\right) = \begin{cases} e^{-t^{\frac{1}{\alpha}}}, & t \geq 0 \\ 0, & t \leq 0 \end{cases} =: \Phi_{\alpha}(t).$$

Das heißt, die Zufallsgröße $M_n/n^{\frac{1}{\alpha}}$ konvergiert in Verteilung gegen Φ_{α} [1].

Beweis: Sei $t > 0$. Mit Gleichung (3.3) erhalten wir

$$P\left(\frac{M_n}{n^{\frac{1}{\alpha}}} \leq t\right) = P(M_n \leq n^{\frac{1}{\alpha}}t) = F^n(n^{\frac{1}{\alpha}}t).$$

Die Zufallsgrößen X_i sind Pareto-verteilt mit Parameter $\alpha > 0 \forall i$ und $n^{\frac{1}{\alpha}}t > 1$ für genügend großes n . Daraus folgt

$$P\left(\frac{M_n}{n^{\frac{1}{\alpha}}} \leq t\right) = \left(1 + \frac{-1}{(n^{\frac{1}{\alpha}}t)^{\alpha}}\right)^n = \left(1 + \frac{-1}{nt^{\alpha}}\right)^n \rightarrow e^{-t^{\frac{1}{\alpha}}}, n \rightarrow \infty. \quad (3.7)$$

Wenn $t < 0$, folgt $P\left(\frac{M_n}{n^{\frac{1}{\alpha}}} \leq t\right) = 0$. □

Definition 3.11 Die Verteilungsfunktion Φ_{α} heißt **Frechet-Verteilung** mit Parameter $\alpha > 0$ [1].

Wir haben nun gezeigt, dass die Pareto-Verteilung gemäß Satz 3.10 im Max-Anziehungsbereich der Frechet-Verteilung liegt.

Satz 3.12 Die Verteilungsfunktion F von X_i habe die Form

$$F(t) = \begin{cases} 0, & t \leq 0, \\ 1 - (1 - t)^{\alpha}, & t \in [0, 1], \\ 1, & t \geq 1, \end{cases}$$

wobei $\alpha > 0$ ein Parameter ist. Dann gilt:

$$\lim_{n \rightarrow \infty} P\left(n^{\frac{1}{\alpha}}(M_n - 1) \leq t\right) = \begin{cases} e^{-(-t)^{\alpha}}, & t \leq 0 \\ 1, & t \geq 0 \end{cases} =: \Psi_{\alpha}(t).$$

Das heißt, die Zufallsgröße $n^{\frac{1}{\alpha}}(M_n - 1)$ konvergiert in Verteilung gegen Ψ_{α} [1].

Beweis: Sei $t \leq 0$. Wir verwenden wieder die Gleichung (3.3) und erhalten

$$P\left(n^{\frac{1}{\alpha}}(M_n - 1) \leq t\right) = P\left(M_n \leq 1 + tn^{\frac{1}{\alpha}}\right) = F^n\left(1 + tn^{\frac{1}{\alpha}}\right).$$

Da $0 \leq 1 + tn^{\frac{1}{\alpha}} \leq 1$ für genügend großes n , können wir die Vorschrift der Verteilungs-

funktion F anwenden. Dabei ergibt sich

$$P\left(n^{\frac{1}{\alpha}}(M_n - 1) \leq t\right) = \left(1 - \left(1 - \left(1 + tn^{\frac{1}{\alpha}}\right)\right)^{\alpha}\right)^n = \left(1 + \frac{-(-t)^{\alpha}}{n}\right)^n \rightarrow e^{-(-t)^{\alpha}}, n \rightarrow \infty.$$

Für $t \geq 0$ ist $1 + tn^{\frac{1}{\alpha}} \geq 1$, sodass $F(1 + tn^{\frac{1}{\alpha}}) = 1$ und deshalb auch $P\left(n^{\frac{1}{\alpha}}(M_n - 1) \leq t\right) = 1$. \square

Definition 3.13 Die Verteilungsfunktion Ψ_{α} heißt **Weibull-Verteilung** mit Parameter $\alpha > 0$ [1].

Das Resultat dieses Beispiels ist, dass die Verteilungsfunktion F gemäß Satz 3.12 im Max-Anziehungsbereich der Weibull-Verteilung liegt. Nun haben wir drei verschiedene Extremwertverteilungen hergeleitet, deren Zufallsgrößen unterschiedliche Werte annehmen können. Gumbel-verteilte Zufallsgrößen können sowohl positive als auch negative Werte annehmen. Dagegen nehmen Weibull-verteilte Zufallsgrößen nur negative und Frechet-verteilte Zufallsgrößen nur positive Werte an. Im späteren Verlauf der Arbeit werden wir sehen, dass sich jede mögliche Extremwertverteilung auf eine der drei bekannten Extremwertverteilungen zurückführen lässt. Dieses Resultat ist der Satz von Fisher-Tippett.

Im folgenden Kapitel werden wir einen kleinen Einblick in die Max-Anziehungsbereiche der einzelnen Extremwertverteilungen geben. Wir sind dann in der Lage Verteilungsfunktionen zu identifizieren, die für $n \rightarrow \infty$ gegen eine der drei Extremwertverteilungen konvergieren.

3.2 Max-Anziehungsbereiche

Wir erinnern uns an die Herleitung der Frechet-Verteilung. Dort konnte man zeigen, dass die Verteilungsfunktion mit der Tailfunktion $t^{-\alpha}$, $t \geq 1$ im Max-Anziehungsbereich von Φ_{α} liegt. Als Resultat dieses Unterabschnitts werden wir sehen, dass der Anziehungsbereich von Φ_{α} aus allen Verteilungsfunktionen besteht, deren Tailfunktionen sich ähnlich verhalten wie $t^{-\alpha}$. Dazu benötigen wir folgende wichtige Definitionen [15]:

Definition 3.14 Sei A eine σ -Algebra auf $X \neq \emptyset$ und B eine σ -Algebra auf $Y \neq \emptyset$. Eine Abbildung $f: X \leftarrow Y$ heißt A - B -messbar, wenn $f^{-1}(b) \in A \forall b \in B$.

Definition 3.15 Eine messbare Funktion f auf $(0, \infty)$ heißt regulär variierend in $+\infty$ mit Index $\alpha \in \mathbb{R}$, falls

$$\lim_{t \rightarrow +\infty} \frac{f(\lambda t)}{f(t)} = \lambda^{\alpha}, \forall \lambda > 0.$$

Bezeichnung: $f \in RV_{\alpha}$ [1]

Diese Definition bedeutet, dass regulär variierende Funktionen sich asymptotisch wie Potenzfunktionen verhalten.

Beispiel 3.16 Wir betrachten die Funktion $f(t) = ct^\alpha$, wobei $c > 0$. Diese Funktion ist regulär variierend mit Index α , weil

$$\frac{f(\lambda t)}{f(t)} = \frac{c(\lambda t)^\alpha}{ct^\alpha} = \lambda^\alpha, \forall \lambda > 0.$$

Definition 3.17 Eine Funktion f ist langsam variierend in $+\infty$ wenn $f \in RV_0 [1]$, das heißt

$$\lim_{t \rightarrow +\infty} \frac{f(\lambda t)}{f(t)} = 1, \forall \lambda > 0.$$

Das heißt, dass langsam variierende Funktionen asymptotisch konstant sind.

Beispiel 3.18 Die Funktion $f(t) = (\log t)^\beta$, mit $\beta \in \mathbb{R}$, ist langsam variierend in $+\infty$, denn

$$\lim_{t \rightarrow +\infty} \frac{(\log \lambda t)^\beta}{(\log t)^\beta} = \lim_{t \rightarrow +\infty} \left(\frac{\log \lambda + \log t}{\log t} \right)^\beta = \lim_{t \rightarrow +\infty} \left(\frac{\log \lambda}{\log t} + 1 \right)^\beta = 1, \forall \lambda > 0.$$

Nun können wir eine Aussage treffen, welche Eigenschaften eine Verteilungsfunktion F haben muss, damit sie im Max-Anziehungsbereich der Frechet-Verteilung liegt [1].

Satz 3.19 Eine Verteilungsfunktion F mit rechtem Endpunkt x^* liegt im Max-Anziehungsbereich der Frechet-Verteilung Φ_α mit $\alpha > 0$, genau dann, wenn die folgenden zwei Bedingungen erfüllt sind:

$$1. x^* = +\infty$$

2. Die Tailfunktion \bar{F} ist regulär variierend mit Index $-\alpha$, das heißt

$$\lim_{t \rightarrow +\infty} \frac{1 - F(\lambda t)}{1 - F(t)} = \lambda^{-\alpha}, \forall \lambda > 0.$$

Sind die beiden Bedingungen erfüllt, so kann man zeigen, dass

$$\frac{\max\{X_1, \dots, X_n\}}{a_n} \xrightarrow{d} \Phi_\alpha, n \rightarrow \infty,$$

$$\text{mit } a_n = F^{-1} \left(1 - \frac{1}{n} \right).$$

Zum Beweis des Satzes 3.19 benötigen wir noch ein Lemma [14].

Lemma 3.20 Sei $x^* = +\infty$ und $\bar{F} \in RV_{-\alpha}$. Dann gilt $\lim_{n \rightarrow \infty} n\bar{F}(a_n) = 1$.

Beweis: Nachzulesen in [14]. □

Der Beweis des Satzes 3.19 ist sehr kurz und übersichtlich, so dass wir ihn von der Idee her kurz erläutern wollen.

Beweis von Satz 3.19 [14:] Wir beschränken uns nur auf die Rückrichtung und nehmen somit an, dass $x = +\infty$ und $\bar{F} \in RV_{-\alpha}$. Sei $t > 0$. Nach dem eben genannten Lemma gilt $n\bar{F}(a_n) \rightarrow 1$. Da \bar{F} regulär variierend mit Index $-\alpha$ ist, ergibt sich

$$n\bar{F}(ta_n) = n\bar{F}(a_n) \frac{\bar{F}(ta_n)}{\bar{F}(a_n)} \rightarrow 1 \cdot t^{-\alpha} = t^{-\alpha}, \quad n \rightarrow \infty.$$

Mit $\bar{F}(ta_n) = \frac{t^{-\alpha}}{n}$ folgt schließlich

$$P\left(\frac{M_n}{a_n} \leq t\right) = (1 - \bar{F}(ta_n))^n \rightarrow e^{-t^{-\alpha}} = \Phi_{\alpha}(t), \quad n \rightarrow \infty,$$

wodurch $\frac{M_n}{a_n} \xrightarrow{d} \Phi_{\alpha}, n \rightarrow \infty$ bewiesen ist. □

Verteilungen, die nun im Max-Anziehungsbereich der Frechet-Verteilung liegen, sind zum Beispiel die Pareto-Verteilung, die Cauchy-Verteilung und die Burr-Verteilung [14]. Es kann ebenfalls die Max-Anziehungsbereiche der Gumbel- und Weibull-Verteilung hergeleitet werden. Diese beiden Fälle werden in dieser Arbeit jedoch nur angedeutet.

Der Max-Anziehungsbereich der Weibull-Verteilung ähnelt dem der Frechet-Verteilung. Der einzige Unterschied ist dabei, dass der rechte Endpunkt x^* der Weibull-Verteilung endlich ist, was eine etwas andere Form des Variationsbegriffes fordert [14].

Definition 3.21 Eine messbare Funktion f auf $(0, \infty)$ heißt regulär variierend in 0 mit Index $\alpha \in \mathbb{R}$, falls

$$\lim_{t \rightarrow 0} \frac{f(\lambda t)}{f(t)} = \lambda^{\alpha}, \quad \alpha > 0.$$

Bezeichnung: $f \in RV_{\alpha}(0)$.

Beispiel 3.22 Die Funktion $f(t) = t^{\alpha}$ ist regulär variierend in 0 mit Index $\alpha \in \mathbb{R}$.

Damit können wir nun den Max-Anziehungsbereich der Weibull-Verteilung beschreiben [1].

Satz 3.23 Eine Verteilungsfunktion F mit rechtem Endpunkt x^* liegt im Max-Anziehungsbereich der Weibull-Verteilung Ψ_α genau dann, wenn die folgenden zwei Bedingungen erfüllt sind:

1. $x^* < \infty$

2. Die Funktion $1 - F(x^* - t)$ ist regulär variierend in 0 mit Index α , das heißt:

$$\lim_{t \downarrow 0} \frac{1 - F(x^* - \lambda t)}{1 - F(x^* - t)} = \lambda^\alpha, \quad \forall \lambda > 0.$$

Dieser Satz besagt, dass die Verteilungen, die im Max-Anziehungsbereich der Weibull-Verteilung liegen, am rechten Rand ein Potenzverhalten aufweisen, z.B.: $\bar{F}(t) = k(x^* - t)^\alpha$.

Zuletzt betrachten wir nun noch den Max-Anziehungsbereich der Gumbel-Verteilung G . Dieser lässt sich etwas einfacher charakterisieren als bei Frechet- und Weibull-Verteilungen [1].

Satz 3.24 Eine Verteilungsfunktion F mit rechtem Endpunkt x^* liegt im Max-Anziehungsbereich der Gumbel-Verteilung $G(t) = e^{e^{-t}}$ genau dann, wenn es eine positive messbare Funktion $g(t)$ gibt, sodass

$$\lim_{t \uparrow x^*} \frac{\bar{F}(t + xg(t))}{\bar{F}(t)} = e^{-x}, \quad \forall x \in \mathbb{R}.$$

Dabei kann x^* diesmal endlich oder unendlich sein.

Verteilungen, die im Max-Anziehungsbereich der Gumbel-Verteilung liegen, sind zum Beispiel die Exponentialverteilung (siehe Satz 3.8), die Normalverteilung und die Loggamma-Verteilung [14]. Im folgenden Abschnitt wollen wir zeigen, dass sich jede beliebige Extremwertverteilung auf eine der drei bekannten Extremwertverteilungen zurückführen lässt.

3.3 Satz von Fisher-Tippett

Wir haben nun drei verschiedene Typen von Extremwertverteilungen hergeleitet und gesehen, dass verschiedene Verteilungen gegen diese bekannten Extremwertverteilungen konvergieren, falls $n \rightarrow \infty$. Wir werden nun sehen, dass sich jede beliebige Extremwertverteilung auf eine der drei definierten Verteilungen zurückführen lässt. Dazu

gibt uns der Satz von Fisher-Tippett Auskunft [14].

Satz 3.25 (Satz von Fisher-Tippett) *Eine Verteilung G ist eine Extremwertverteilung genau dann, wenn es $c > 0$, $d \in \mathbb{R}$ und $\gamma \in \mathbb{R}$ gibt mit*

$$G(t) = G_\gamma(ct + d).$$

Dabei ist G_γ die verallgemeinerte Extremwertverteilung, die durch $G_\gamma(t) = \exp \left\{ -(1 + \gamma t)^{\frac{-1}{\gamma}} \right\}$ für $1 + \gamma t > 0$ definiert ist.

Wenn man die Funktion G_γ etwas näher betrachtet, kann Folgendes festgestellt werden (siehe Abbildung 3.3 und 3.4):

1. Für $\gamma > 0$ ist G_γ vom gleichen Typ wie die Frechet-Verteilung $\Phi_{\frac{1}{\gamma}}(t) = e^{(-t)^{\frac{1}{\gamma}}}$, $t > 0$.
2. Für $\gamma < 0$ ist G_γ vom gleichen Typ wie die Weibull-Verteilung $\Psi_{\frac{-1}{\gamma}}(t) = e^{-(-t)^{\frac{1}{\gamma}}}$, $t < 0$.
3. Für $\gamma = 0$ ist G_γ nicht definiert, weshalb wir $\gamma \rightarrow 0$ betrachten:

$$\lim_{\gamma \rightarrow 0} (1 + \gamma t)^{\frac{-1}{\gamma}} = e^{-t}.$$

Somit ist dann $G_0(t) = e^{-e^{-t}}$, $t \in \mathbb{R}$ die Gumbel-Verteilung.

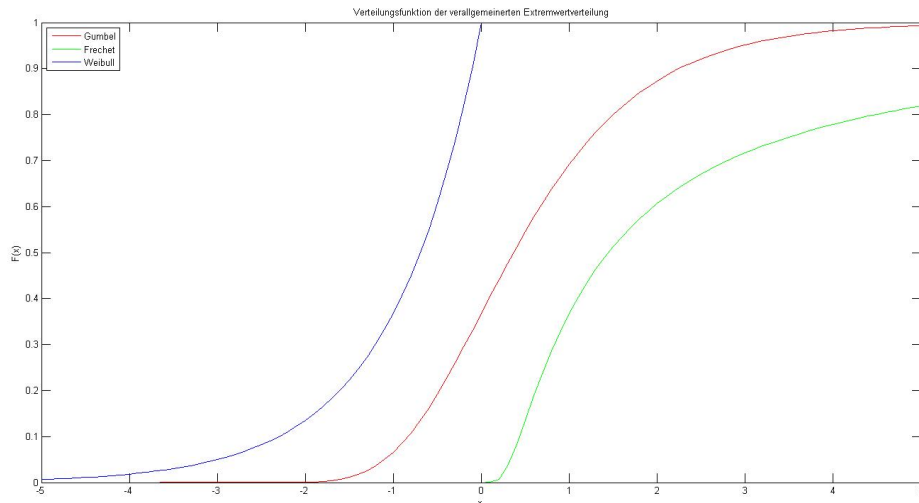


Abbildung 3.3: Graphen der drei Extremwertverteilungen, $\gamma = 0$ für Gumbel, $\gamma = -1$ für Weibull und $\gamma = 1$ für Frechet

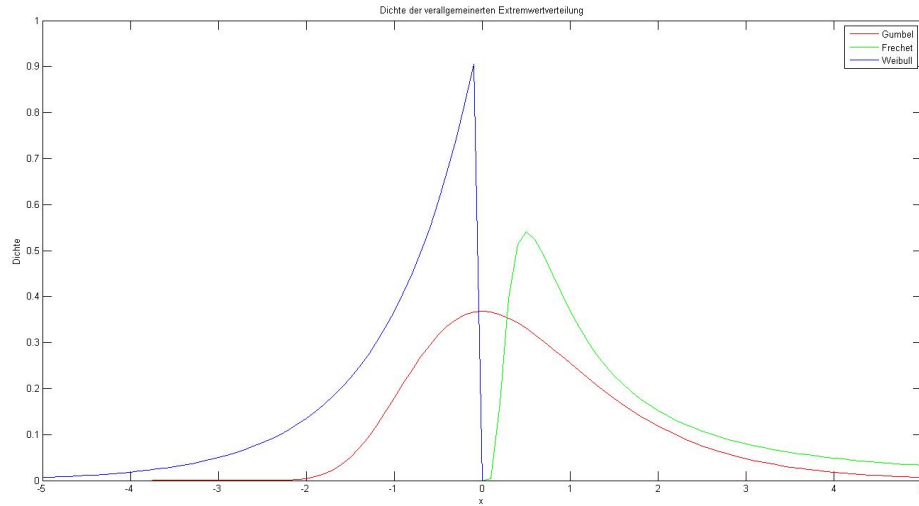


Abbildung 3.4: Graphen der drei Extremwertverteilungen, $\gamma = 0$ für Gumbel, $\gamma = -1$ für Weibull und $\gamma = 1$ für Frechet

Man kann die Verallgemeinerte Extremwertverteilung G_γ ebenfalls für standardisierte Zufallsgrößen definieren. Dazu muss nur $t = \frac{x-\mu}{\sigma}$ gesetzt werden. Somit erhalten wir

$$G_{\gamma,\mu,\sigma}(x) = \exp \left\{ - \left(1 + \gamma \frac{x-\mu}{\sigma} \right)^{1/\gamma} \right\}, \quad 1 + \gamma \frac{x-\mu}{\sigma} > 0. \quad (3.8)$$

Hierbei ist $\gamma \in \mathbb{R}$ der Formparameter, $\mu \in \mathbb{R}$ der Lageparameter und $\sigma > 0$ der Skalierungsparameter. Diese Gestalt der Extremwertverteilung werden wir im späteren Verlauf des Kapitels benutzen, um die Parameter γ , μ und σ zu schätzen. Vorher soll jedoch dem Hauptsatz der Theorie der Blockmaxima nochmals Aufmerksamkeit geschenkt werden.

Der Satz von Fisher-Tippett ist für die Theorie der Blockmaxima sehr wichtig, da er alle drei bekannten Extremwertverteilungen in einer allgemeinen Formel vereinigt und ebenfalls besagt, dass sich jede beliebige Extremwertverteilung in dieser allgemeinen Form darstellen lässt. Er ist somit das Pendant zum zentralen Grenzwertsatz für Summen und beantwortet uns die Frage zur Gleichung (3.5). Das Resultat des Satzes von Fisher-Tippett bringt uns in der Anwendung den Vorteil, dass wir nicht entscheiden müssen, welche der drei Extremwertverteilungen benutzt werden soll. Die Festlegung auf eine der drei bekannten Typen folgt durch die Schätzung der Parameter, auf die wir in Kapitel 3.5 eingehen werden.

Um den Satz von Fisher-Tippett beweisen zu können, benötigen wir noch etwas Vorarbeit.

Definition 3.26 Es sei (Ω, E, P) ein Wahrscheinlichkeitsraum. Als **degenerierte Zufallsgröße** bezeichnet man jede messbare Abbildung X von (Ω, E) in $(\mathbb{R}, B(\mathbb{R}))$ mit der Eigenschaft, dass

$$X(s) = c, \text{ bzw. } P(X = c) = 1$$

für fast alle $s \in \Omega$ und $c \in \mathbb{R}$. Die zugehörige Verteilungsfunktion F von X nennt man dann **degenerierte Verteilungsfunktion** [2].

Eine degenerierte Verteilungsfunktion ist somit ein Einpunktmaß, das heißt, das zugrunde liegende Zufallsexperiment hat nur einen möglichen Ausgang.

Definition 3.27 Eine nicht degenerierte Verteilungsfunktion heißt **max-stabil** [14], falls es für alle natürlichen Zahlen n Konstanten $c_n > 0$ und $d_n \in \mathbb{R}$ gibt mit

$$G^n(c_n t + d_n) = G(t).$$

Beispiel 3.28 Die Gumbel-Verteilung $G(t) = e^{-e^{-t}}$ ist max-stabil, denn

$$G^n(t + \log n) = e^{-ne^{-(t+\log n)}} = e^{-ne^{-t} e^{-\log n}} = e^{-\frac{ne^{-t}}{n}} = e^{-e^{-t}}.$$

Analog kann man das auch für die Frechet- und Weibull-Verteilung zeigen.

Satz 3.29 Eine Verteilungsfunktion G ist max-stabil, genau dann, wenn G eine Extremwertverteilung ist.

Beweis: Siehe [14]. □

Satz 3.30 Eine max-stabile Verteilung G ist vom selben Typ wie eine der folgenden Verteilungen: Gumbel Λ , Frechet Φ_α mit $\alpha > 0$ oder Weibull Ψ_α mit $\alpha > 0$.

Dieser Satz zusammen mit Satz 3.29 impliziert den Satz von Fisher-Tippett, dessen Beweiskizze wir kurz vorstellen wollen.

Beweis: (Satz von Fisher-Tippett) Im Folgenden beweisen wir nur die Hinrichtung des Satzes 3.30 und gehen somit davon aus, dass G max-stabil ist. Wir wollen dann darauf schließen, dass G vom gleichen Typ ist wie eine der drei bekannten Extremwertverteilungen. Man kann beweisen, dass es Konstanten $c(s) > 0$ und $d(s) \in \mathbb{R}$ gibt mit

$$G^s(c(s)t + d(s)) = G(t) \tag{3.9}$$

falls G max-stabil ist [14]. Dies ist der Ausgangspunkt des Beweises. Die Idee ist nun, durch geschicktes Umstellen nach $G(t)$ auf die allgemeine Extremwertverteilung zu stoßen. Zunächst wird die Gleichung (3.9) etwas umgeformt, so dass wir folgende Form erhalten:

$$-\log(-\log G(c(s)t + d(s))) - \log s = -\log(-\log G(t)). \quad (3.10)$$

Wir setzen

$$\psi(t) = -\log(-\log G(t)) \quad (3.11)$$

und erhalten

$$\psi(c(s)t + d(s)) - \log s = \psi(t).$$

Da G unbekannt ist, ist auch $\psi(t)$ unbekannt. Wenn man also $\psi(t)$ berechnen kann, so kann dann auf G geschlossen werden. Wir betrachten $U(y) = \psi^{-1}(t) = \inf\{t \in \mathbb{R} \mid \psi(t) \geq y\}$ als die verallgemeinerte Inverse von $\psi(t)$. Ziel ist es, für $U(y)$ einen geschlossenen Ausdruck zu finden. Dann kann man durch die Umkehrfunktion auf $\psi(t)$ schließen und somit auf G . Durch geschicktes Umstellen (siehe [14]) kann man erreichen, dass

$$U(y) - U(0) = \frac{U(y + \log s) - U(\log s)}{c(s)}.$$

Setzt man anschließend noch $\log s = z$, so folgt:

$$U(y) - U(0) = \frac{U(y + z) - U(z)}{c(e^z)}.$$

Mit dieser hergeleiteten Gleichung lässt sich nun die Behauptung zeigen. Dafür unterscheidet man zwei Fälle:

1. $c(e^z) \equiv 1$ und
2. $c(e^z) \neq 1$ für mindestens ein z .

Im ersten Fall kann nachgewiesen werden [14], dass $U(y) = p \cdot y + b$ mit $b = U(0)$. Man erhält nun die Funktion $\psi(t)$, indem von $t = p \cdot y + b$ die Umkehrfunktion gebildet wird. Wir erhalten also

$$\psi(t) = \frac{t - b}{p}. \quad (3.12)$$

Jedoch gilt auerdem $\psi(t) = -\log(-\log G(t))$ (3.11), laut Definition. Durch Gleichsetzen von Gleichung (3.12) und Gleichung (3.11) und anschlieend Umstellen nach $G(t)$ folgt:

$$G(t) = e^{-e^{-\psi(t)}} = e^{-e^{-\frac{t-b}{p}}}.$$

$G(t)$ ist dabei vom gleichen Typ wie die Gumbel-Verteilung.

Im zweiten Fall wird hnlich argumentiert, wie im ersten. Man kann zeigen [14], dass

$$U(y) = \alpha + \beta e^{py}.$$

Zum Schluss berechnen wir wieder $\psi(t)$, indem die Umkehrfunktion von $t = \alpha + \beta e^{py}$ gebildet wird. Wir erhalten somit

$$\psi(t) = \frac{1}{p} \log\left(\frac{z - \alpha}{\beta}\right), \quad (3.13)$$

fr $z > \alpha$, falls $\beta > 0$, bzw. fr $z < \alpha$, falls $\beta < 0$. Durch Umstellen von (I) nach $G(t)$ und anschlieend Einsetzen von Gleichung (3.13) folgt:

$$G(t) = e^{-e^{-\psi(t)}} = \exp\left\{-\left(\frac{z - \alpha}{\beta}\right)^{\frac{-1}{p}}\right\}.$$

Ist nun $z > \alpha$, falls $\beta > 0$, dann ist $G(t)$ vom gleichen Typ wie die Frechet-Verteilung. Wenn $z < \alpha$, falls $\beta < 0$, dann ist $G(t)$ vom selben Typ wie die Weibull-Verteilung, womit der Satz von Fisher-Tippett bewiesen ist.

□

Erwhnenswert ist der Aspekt, dass die gerade bewiesenen Erkenntnisse auch fr Minima gelten [1]. Dabei betrachte man wieder unabhngig, identisch verteilte Zufallsgren deren Verteilung die folgende Eigenschaft aufweist:

$$F(x) = 1 - F(-x), \quad \forall x.$$

Benutzt man die Tatsache, dass

$$\min \{X_1, \dots, X_n\} = -\max \{-X_1, \dots, -X_n\}, \quad (3.14)$$

kann nun $M_n^* = \max \{-X_1, \dots, -X_n\}$ gesetzt werden. Daraus folgt:

$$\lim_{n \rightarrow \infty} P \left(\frac{M_n^* - b_n}{a_n} \leq t \right) = G_\gamma(t).$$

Wenden wir nun darauf Gleichung (3.14) an, folgt somit:

$$\lim_{n \rightarrow \infty} P \left(\frac{-\min \{X_1, \dots, X_n\} - b_n}{a_n} \leq t \right) = \lim_{n \rightarrow \infty} P \left(\frac{\min \{X_1, \dots, X_n\} + b_n}{a_n} \geq -t \right)$$

und schließlich:

$$\lim_{n \rightarrow \infty} P \left(\frac{\min \{X_1, \dots, X_n\} + b_n}{a_n} \geq -t \right) = \lim_{n \rightarrow \infty} 1 - P \left(\frac{\min \{X_1, \dots, X_n\} + b_n}{a_n} \leq -t \right) = 1 - G_\gamma(-t).$$

3.4 Extremwertverhalten von stationären Zeitreihen

Wie wir wissen, ist die Voraussetzung des Satzes von Fisher-Tippett, dass die einzelnen Realisierungen der Blöcke identisch verteilt und stochastisch unabhängig sein müssen. Da uns aber das Extremalverhalten von stationären Zeitreihen interessiert, können wir die Bedingung der Unabhängigkeit nicht erfüllen. Die Frage ist nun, wie sich die möglichen Korrelationen auf das Extremalverhalten von stationären Zeitreihen auswirken. Unsere Vermutung ist, dass eine schwache Abhängigkeitsstruktur das Extremwertverhalten nicht beeinflusst. Wir müssen nun also ein Resultat finden, welches uns erlaubt, die Theorie der Blockmaxima für identisch, unabhängige Zufallsgrößen auch auf stationäre Folgen von Zufallsgrößen anzuwenden, solange die Abhängigkeitsstruktur nicht zu stark ist. Diese noch sehr ungenaue Formulierung werden wir in diesem Kapitel präzisieren.

Unter stationären Folgen können ebenfalls konstante Folgen gesehen werden. Diese Art von stationären Folgen wollen wir jedoch im weiteren Verlauf dieses Kapitels ausschließen, da sie für unser Ziel (Extremwertverhalten von stationären Zeitreihen) keine Bedeutung haben.

Im vorherigen Kapitel haben wir durch das Theorem von Fisher-Tippett gesehen, wenn es Folgen $(a_n) > 0$ und $(b_n) \in \mathbb{R}$ gibt und F im Max-Anziehungsbereich von G liegt,

dann haben die Maxima M_n das folgende Grenzverhalten:

$$\lim_{n \rightarrow \infty} P\left(\frac{M_n - b_n}{a_n} \leq x\right) = \lim_{n \rightarrow \infty} F^n(a_n x + b_n) = G_\gamma(x),$$

wobei G_γ eine verallgemeinerte Extremwertverteilung ist. Bildet man auf beiden Seiten den Logarithmus, so ergibt sich

$$\lim_{n \rightarrow \infty} n(1 - F(a_n x + b_n)) = -\log G_\gamma, \quad (3.15)$$

wobei man beachte, dass $-\log(1 - x) \approx x$ gilt für $x \rightarrow 0$, wenn man den Logarithmus um 1 entwickelt [4]. Wir betrachten nun eine allgemeinere Aussage.

Satz 3.31 *Sei (u_n) eine Folge reeller Zahlen, F eine Verteilungsfunktion und M_n die Blockmaxima. Für $0 \leq \tau \leq \infty$ ist äquivalent:*

- (i) $\lim_{n \rightarrow \infty} n(1 - F(u_n)) = \tau$
- (ii) $\lim_{n \rightarrow \infty} P(M_n \leq u_n) = e^{-\tau}.$

Beweis: Siehe [4]. □

Wir wollen nun einen ähnlichen Grenzwertsatz für stationäre Zeitreihen aufstellen, wie es ihn auch für unabhängige, identisch verteilte Zufallsgrößen gibt. Man kann zeigen, dass es einen solchen Grenzwertsatz gibt, der folgende Bedingungen fordert:

Sei (X_n) eine stationäre Folge von Zufallsgrößen und von einer Realisierung X_i ist F die Randverteilung. Dann wird mit

$$M_n = \max\{X_1, \dots, X_n\}$$

das Maximum der ersten n Zufallsgrößen bezeichnet. (\tilde{X}_n) ist eine Folge mit unabhängig identisch verteilten Zufallsgrößen und gleicher Randverteilung F und es gilt

$$\tilde{M}_n = \max\{\tilde{X}_1, \dots, \tilde{X}_n\}.$$

Wenn die folgenden zwei Bedingungen erfüllt sind, haben die Folgen (X_n) und (\tilde{X}_n) dasselbe Grenzwertverhalten:

1. Die stationäre Folge darf nur eine schwache lang andauernde Abhängigkeit haben.
2. Die Folge darf keine Tendenz aufweisen, dass große Werte in Clustern auftreten.

Bei Bedingung 1 handelt es sich um die sogenannte „Mischungseigenschaft“ [4]. Formal kann sie folgendermaßen definiert werden:

Definition 3.32 Eine stationäre Folge X_1, X_2, \dots erfüllt die **Mischungsbedingung** $D(u_n)$ für eine Folge (u_n) , falls für alle $i_1 < \dots < i_p < j_1 < \dots < j_q$ mit $j_1 - i_p > l$ gilt:

$$|P(X_{i_1} \leq u_n, \dots, X_{i_p} \leq u_n, X_{j_1} \leq u_n, \dots, X_{j_q} \leq u_n) - P(X_{i_1} \leq u_n, \dots, X_{i_p} \leq u_n) \cdot P(X_{j_1} \leq u_n, \dots, X_{j_q} \leq u_n)| \leq g(n, l),$$

wobei $g(n, l_n) \rightarrow 0$, für eine Folge l_n , so dass $\frac{l_n}{n} \rightarrow 0$, wenn $n \rightarrow \infty$.

Man kann beobachten, dass wenn die Zufallsgrößen unabhängig sind, der obige Ausdruck 0 ist. Für uns ist jedoch der Fall interessant, dass die Zufallsgrößen nicht unabhängig sind. Wenn die Mischungsbedingung $D(u_n)$ gilt und die X_{i_1}, \dots, X_{i_p} weit von den X_{j_1}, \dots, X_{j_q} entfernt sind, so dass die Differenz der Wahrscheinlichkeiten gegen 0 geht, so hat dies keinen Einfluss auf das Grenzverhalten der Maxima. Dieses Resultat wird im folgenden Satz beschrieben:

Satz 3.33 Sei (X_n) eine stationäre Folge von Zufallsgrößen. Wenn es Folgen von $(a_n) > 0$ und $(b_n) \in \mathbb{R}$ gibt, so dass

$$P\left(\frac{M_n - b_n}{a_n} \leq x\right) \rightarrow G(x)$$

gilt, wobei G eine nicht-entartete Verteilungsfunktion ist und die Bedingung $D(u_n)$ ist erfüllt mit $u_n = a_n x + b_n$ für alle $x \in \mathbb{R}$. Dann ist G eine Extremwertverteilung.

Beweis: Nachzulesen in [4]. □

Wir wissen nun, dass die skalierten Maxima einer stationären Folge von Zufallsgrößen bei erfüllter $D(u_n)$ -Bedingung gegen eine Extremwertverteilung konvergieren. Doch ist es noch unklar, wenn $P\left(\frac{M_n - b_n}{a_n} \leq x\right) \rightarrow G(x)$ und $P\left(\frac{\tilde{M}_n - b_n}{a_n} \leq x\right) \rightarrow H(x)$ gilt, dass dann auch $G(x) \equiv H(x)$ gilt. Um dieses Problem zu lösen, benötigt man noch die zweite Bedingung, die wir kürzlich eingeführt haben. Es handelt sich hierbei um die „Anti-Cluster-Bedingung“, welche wie folgt definiert ist [4]:

Definition 3.34 Sei (X_n) eine Folge von stationären Zufallsgrößen und (u_n) eine Folge

in \mathbb{R} . Die **Anti-Cluster-Bedingung** $D'(u_n)$ ist für (X_n) erfüllt, falls gilt

$$\lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} n \sum_{j=2}^{[n/k]} P(X_1 > u_n, X_j > u_n) = 0.$$

Diese Bedingung sorgt dafür, dass Extremwerte nicht in Clustern auftreten, das heißt, dass in der Umgebung von Extremwerten keine weiteren Werte existieren, die dem Maximum sehr nahe kommen. Bildlich kann man sich das an folgendem Diagramm verdeutlichen:

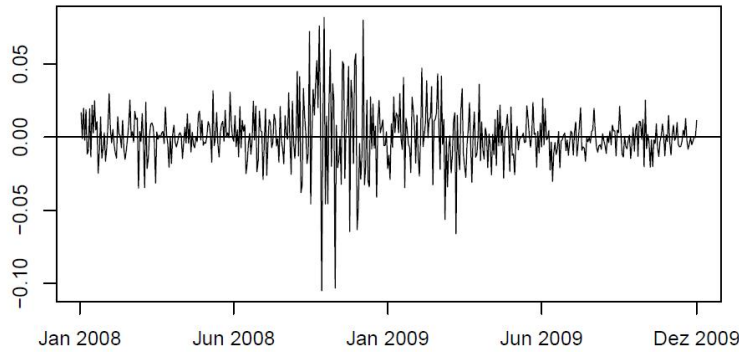


Abbildung 3.5: Die täglichen negativen Log-Renditen des Dow Jones Indexes von Januar 2008 bis Dezember 2009 [4]

Man kann in Abbildung 3.5 erkennen, dass sich zwischen Juni 2008 und Januar 2009 eine extreme Kursentwicklung ergibt. Diese hohen Kurswerte treten alle in einem Cluster auf. Genau dieses Verhalten wollen wir in unseren stationären Zeitreihen ausschließen.

Nun kann folgender Satz formuliert werden:

Satz 3.35 Sei (X_n) eine Folge von stationären Zufallsgrößen mit Verteilungsfunktion F , den Blockmaxima M_n und (u_n) eine Folge. (X_n) erfüllt die Mischungsbedingung $D(u_n)$, die Anti-Cluster-Bedingung $D'(u_n)$ und es gibt ein $0 \leq \tau \leq \infty$. Dann ist äquivalent:

- (i) $\lim_{n \rightarrow \infty} n(1 - F(u_n)) = \tau$
- (ii) $\lim_{n \rightarrow \infty} P(M_n \leq u_n) = e^{-\tau}.$

Beweis: Nachzulesen in [6] und [3]. □

Offensichtlich gilt dieser Satz ebenso für Folgen von stationären Zufallsgrößen. Nun können wir das Hauptresultat dieses Abschnitts formulieren.

Satz 3.36 Sei (X_n) eine stationäre Folge mit Verteilungsfunktion F , wobei F im Max-Anziehungsbereich von G_γ liegt und G_γ eine verallgemeinerte Extremwertverteilung ist.

Das heißt es existieren Folgen $(a_n) > 0$ und $(b_n) \in \mathbb{R}$, so dass gilt

$$\lim_{n \rightarrow \infty} n(1 - F(a_n x + b_n)) = -\log G_\gamma(x), \quad x \in \mathbb{R}. \quad (3.16)$$

Die Folge $(u_n) = (a_n x + b_n)$ erfüllt die Bedingungen $D(u_n)$ und $D'(u_n)$ für alle $x \in \mathbb{R}$. Dann ist Gleichung (3.16) äquivalent zu jeder der folgenden Aussage:

$$\begin{aligned} \lim_{n \rightarrow \infty} P\left(\frac{M_n - b_n}{a_n} \leq x\right) &\rightarrow G_\gamma(x) \\ \lim_{n \rightarrow \infty} P\left(\frac{\tilde{M}_n - b_n}{a_n} \leq x\right) &\rightarrow G_\gamma(x) \end{aligned} \quad (3.17)$$

Beweis: Siehe [3]. □

Somit können wir die Verfahren der Extremwerttheorie für unabhängige, identisch verteilte Zufallsgrößen ebenfalls für stationäre Folgen von Zufallsgrößen anwenden, da das Extremwertverhalten in beiden Fällen identisch ist. Außerdem sind die Mischungseigenschaft und die Anti-Cluster-Bedingung auch nur hinreichende Bedingungen, deren Nichterfülltheit noch lange nicht das Scheitern der Theorie der Blockmaxima impliziert. An dieser Stelle sei erwähnt, dass es ebenfalls Ansätze gibt, für den Fall, dass die Maxima in Clustern auftreten und die Mischungsbedingung ebenfalls nicht erfüllt ist. Hierfür existiert die Theorie des extremalen Index für Blockmaxima ([3] und [4]), welche uns ein Maß für die Abhängigkeit beschreibt, die zwischen den Datenpunkten besteht. Weisen die Daten einen Trend auf, das heißt, nehmen die Maxima im Laufe der Zeit zu oder ab, kann dieser Trend bei der Schätzung der Parameter berücksichtigt werden (nachzulesen in [14]). Da in dieser Arbeit jedoch nur Spotmarktpreise betrachtet werden, deren Zeitreihen stationär und clusterfrei sind, werden die eben genannten Ansätze nicht weiter verfolgt.

3.5 Schätzung der Parameter und des Value-at-Risk

Für die Anwendung der Blockmaxima auf praxisrelevante Daten ist die Schätzung der Parameter ein wichtiger Aspekt, der in diesem Abschnitt genauer beschrieben werden soll. Unser Problem ist es, den Parametervektor $\Theta = (\gamma, \mu, \sigma) \in \mathbb{R}^3$ zu schätzen, so dass die Gestalt der verallgemeinerten Extremwertverteilung für deren Realisierungen X_1, \dots, X_n beschrieben werden kann. Für diese Aufgabe bietet sich die Maximum-Likelihood-Methode (ML) an, die wir im Folgenden detaillierter betrachten wollen. Dafür wird zunächst die Dichte $f_\Theta(x)$ der verallgemeinerten Extremwertverteilung benötigt. Diese erhalten wir, indem von $G_\Theta(x)$ die erste Ableitung gebildet wird:

$$f_{\Theta}(x) = f_{\gamma, \mu, \sigma}(x) = \begin{cases} \frac{1}{\sigma} \left(1 + \gamma \frac{x - \mu}{\sigma}\right)^{-\frac{1}{\gamma} - 1} \exp\left\{-\left(1 + \gamma \frac{x - \mu}{\sigma}\right)^{1/\gamma}\right\}, & 1 + \gamma \frac{x - \mu}{\sigma} > 0 \\ 0, & \text{sonst} \end{cases} \quad (3.18)$$

Für $\gamma = 0$ gilt:

$$f_{\Theta}(x) = f_{0, \mu, \sigma}(x) = \frac{1}{\sigma} e^{-\frac{x - \mu}{\sigma}} \exp\left\{-e^{-\frac{x - \mu}{\sigma}}\right\}, \quad x \in \mathbb{R}. \quad (3.19)$$

Mit Hilfe der Dichten sind wir nun in der Lage die Log-Likelihoodfunktion aufzustellen:

$$l(\Theta) := l(\Theta | x_1, \dots, x_n) := \sum_{i=1}^n \log f_{\Theta}(x_i). \quad (3.20)$$

Wir betrachten nun die zwei Fälle $\gamma \neq 0$ und $\gamma = 0$. Für $\gamma \neq 0$ ergibt sich durch Einsetzen der Dichte (3.18) in (3.20)

$$l(\Theta) = -n \log \sigma - \left(\frac{1}{\gamma} + 1\right) \sum_{i=1}^n \log \left(1 + \gamma \frac{x_i - \mu}{\sigma}\right) - \sum_{i=1}^n \left(1 + \gamma \frac{x_i - \mu}{\sigma}\right)^{-\frac{1}{\gamma}},$$

falls $1 + \gamma \frac{x_i - \mu}{\sigma} > 0$ für alle $i = 1, \dots, n$ und $l(\Theta) = -\infty$ sonst. Für $\gamma = 0$ gilt:

$$l(\Theta) = -n \log \sigma - \sum_{i=1}^n \frac{x_i - \mu}{\sigma} - \sum_{i=1}^n e^{-\frac{x_i - \mu}{\sigma}}.$$

Um nun den Maximum-Likelihood-Schätzer berechnen zu können, muss der Gradient von $l(\Theta)$ gleich 0 gesetzt werden:

$$\nabla l = \begin{pmatrix} \frac{\partial l}{\partial \gamma} \\ \frac{\partial l}{\partial \mu} \\ \frac{\partial l}{\partial \sigma} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \quad (3.21)$$

Das entstandene nichtlineare Gleichungssystem (3.21) kann im Allgemeinen nicht exakt gelöst werden. Es muss an dieser Stelle auf numerische Lösungsverfahren wie zum Beispiel Gradientenverfahren zurückgegriffen werden, deren Theorie jedoch in dieser

Arbeit nicht erläutert wird. Die Lösung von (3.21) ist unser gesuchter Parametervektor:

$$\hat{\Theta} = (\hat{\gamma}, \hat{\mu}, \hat{\sigma}) = \operatorname{argmax} l(\gamma, \mu, \sigma). \quad (3.22)$$

Mit Gleichung (3.22) kann nun auch der Value-at-Risk geschätzt werden, den wir im Kapitel 2.5 als das $(1 - \alpha)$ -Quantil der zu Grunde liegenden Verteilung definiert haben. Der VaR ist also der Wert, der mit einer Wahrscheinlichkeit von $1 - \alpha$ in der Betrachtungsperiode nicht überschritten wird. Wir schätzen den Value-at-Risk, indem wir die folgende Gleichung betrachten:

$$G_{\hat{\gamma}, \hat{\mu}, \hat{\sigma}}(VaR_{1-\alpha}) = 1 - \alpha. \quad (3.23)$$

Es kann sein, dass Gleichung (3.23) mehrere Lösungen besitzt. Ist das der Fall, so betrachtet man immer die kleinste. Setzt man in Gleichung (3.23) für $G_{\hat{\gamma}, \hat{\mu}, \hat{\sigma}}$ die geschätzte Verteilungsfunktion ein und stellt anschließend nach $VaR_{1-\alpha}$ um, so erhalten wir [5]:

$$VaR_{1-\alpha} = \begin{cases} \hat{\mu} - \frac{\hat{\sigma}}{\hat{\gamma}} (1 - (-\log(1 - \alpha))^{-\hat{\gamma}}), & \hat{\gamma} \neq 0 \\ \hat{\mu} - \hat{\sigma} \log(-\log(1 - \alpha)), & \hat{\gamma} = 0 \end{cases}. \quad (3.24)$$

Im Falle von $\hat{\gamma} < 0$ (Weibullverteilung) ergibt sich ein interessantes Ergebnis. Wie wir seit den Max-Anziehungsbereichen wissen, besitzt die Weibullverteilung einen rechten Endpunkt $x^* < \infty$. Diese Stelle ist nach Definition des Value at Risk gleich VaR_1 . Es ist also der höchste Wert, der niemals überschritten wird. Den zugehörigen Schätzer erhält man, indem in Gleichung (3.24) $\alpha = 0$ gesetzt wird:

$$VaR_1 = \hat{\mu} - \frac{\hat{\sigma}}{\hat{\gamma}}.$$

Wichtig ist hier zu wissen, dass der hergeleitete $VaR_{1-\alpha}$ für die Dauer eines ganzen Blocks gilt. Dazu bezieht sich $VaR_{1-\alpha}$ auf Maxima und nicht auf alltägliche Preise wie der herkömmliche Value at Risk. Die Größe $VaR_{1-\alpha}$ gibt uns somit an, wie hoch der maximale Strompreis ist, der innerhalb eines Blocks mit einer Wahrscheinlichkeit von $1 - \alpha$ nicht überschritten wird.

Damit ist nun ein wichtiges Problem für die Anwendung auf praxisrelevante Daten gelöst. Es bleibt nun nur noch die Frage zu klären, wie die Wahl der Blocklänge ausfällt.

3.6 Wahl der Blocklänge

Die Wahl der Blocklänge erweist sich als schwieriges Problem, da die Größe eines Blocks den Datenumfang beeinflusst. Wählt man die Blocklänge zu klein, bleiben nach Bestimmung des Maximums viele Daten übrig, die jedoch zum Großteil keine extremen Datenpunkte sind. Eine große Blocklänge würde dieses Problem lösen, filtert aber auch viele große Datenpunkte heraus, die für eine Verteilungsannahme dann unberücksichtigt bleiben. Es muss also nach einer Kompromisslösung gesucht werden.

Ein Ansatz wäre zum Beispiel, das Verfahren der Blockmaxima auf die Zeitreihe anzuwenden und mit einer kleinen Blocklänge zu beginnen. Anschließend misst man die Güte des Ergebnisses, zum Beispiel mit dem Kolmogorov-Smirnov-Test (siehe Anhang B) und lässt sich das p -Level (siehe Anhang B) als Gütemaß ausgeben. Darauf wird die Blocklänge sukzessive erhöht und der Vorgang wiederholt. Dies kann solange fortgeführt werden, bis eine maximale Blocklänge von der Hälfte der Zeitreihenlänge erreicht ist. Mit dieser Länge des Blocks erhält man zwei Datenpunkte, auf deren Grundlage selbstverständlich keine Verteilungsannahme getroffen werden kann. Verwendet wurde der folgende Algorithmus:

Algorithmus 3 optimale Blocklänge

```

Setze optblock = 0
Setze plevel = 0
for  $i = 1$  to Zeitreihenlänge/2 do
  Filtere Daten gemäß Blocklänge  $i$ 
  Passe Daten an Verallgemeinerte Extremwertverteilung an
  ( $\text{Hypothese}, p$ )  $\leftarrow$  Kolmogorov-Smirnov-Test
  if  $p > \text{plevel}$  then
    Setze optblock =  $i$ 
    Setze plevel =  $p$ 
  end if
end for
return (optblock, plevel)

```

Angewendet wurde der Algorithmus auf den Phelix-Index im Base und im Peak. Im Base wurde eine optimale Blocklänge von 286 und im Peak von 284 Tagen berechnet. Bei diesen Blocklängen erhalten wir am Ende einen Datenumfang von jeweils 10 Datenpunkten, was für eine Verteilungsannahme deutlich zu wenig ist. Aus diesem Grund suchen wir nach weiteren lokalen Maxima, deren Blocklängen kleiner ist. Um diese zu finden, wandeln wir Algorithmus 3 so ab, dass jede Blocklänge gespeichert und nach Beendigung der For-Schleife das p -Level in Abhängigkeit der Blocklänge geplottet wird. Im Base und im Peak ergeben sich folgende Graphen:

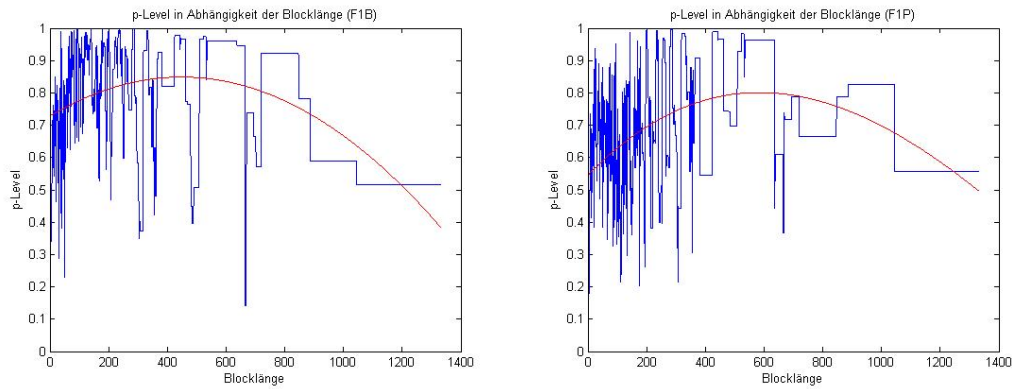


Abbildung 3.6: Verlauf des p -Levels in Abhängigkeit der Blocklänge für Base und Peak mit eingezeichneter Regressionskurve (rot)

Man erkennt in Abbildung 3.6 eine starke Oszillation des p -Levels bei beiden Zeitreihen. Das heißt, dass bei kleiner Variation der Blocklänge das p -Level stark schwankt. Dieses Phänomen lässt sich folgendermaßen erklären:

Innerhalb eines Blocks liegen mehrere größere Datenpunkte, wobei nur einer von ihnen als Maximum identifiziert wird. Alle anderen fallen für die weitere Betrachtung weg. Vergrößert oder verkleinert man nun die Blocklänge, so können große Datenpunkte, die kleiner als das Maximum dieses Blocks sind, in einen neuen Block fallen. Dort besteht wieder die Möglichkeit, dass sie zum Maximum werden. Ist dies der Fall, so erhöht sich die relative Häufigkeit dieses Wertes und beeinflusst somit auch die empirische Verteilungsfunktion. Der maximale Abstand zwischen empirischer und geschätzter Verteilungsfunktion ist die Kenngröße des Kolmogorov-Smirnov-Tests, die wiederum das p -Level in Abhängigkeit des Datenumfangs beeinflusst.

Um dennoch verschiedene Blocklängen zu erhalten, die ein hohes p -Level bewirken, muss Algorithmus 3 nochmals ergänzt werden. Interessant für unser Problem wäre eine Liste von Blocklängen, deren p -Level eine Mindestschranke nicht unterschreitet. Dies lässt sich leicht umsetzen. Für den Phelix-Index im Base erhalten wir unter anderem folgende Blocklängen, die, bezogen auf ein Börsenjahr (250 Tage), eine interessante Interpretationsmöglichkeit mitbringen:

Blocklänge	p -Level	Interpretation
7	0,6559	$\approx 1,5$ Wochen
24	0,8010	≈ 1 Monat

Ähnliche Blocklängen ergeben sich auch für den Phelix-Index im Peak:

Blocklänge	p -Level	Interpretation
7	0,6746	$\approx 1,5$ Wochen
25	0,9380	≈ 1 Monat

Diese Blocklängen erfüllen unsere Kriterien, die wir am Anfang des Abschnitts gestellt haben. Sie sind weder zu groß, so dass nur wenige Daten übrig bleiben, noch sind sie zu klein, so dass die Menge der Maxima aus zu kleinen Datenpunkten besteht, die keine extremen Daten sind. Im folgenden Kapitel werden wir diese Blocklängen bezüglich ihrer resultierenden Quantile und Verteilung untersuchen.

3.7 Anwendung des Verfahrens

In diesem Abschnitt wollen wir nun die Theorie der Blockmaxima auf Zeitreihen anwenden. Zunächst werden autoregressive Zeitreihen betrachtet, um zu sehen wie sich das Verfahren verhält, wenn die Stationarität, die Clusterfreiheit, eine schwache Abhängigkeitsstruktur und zudem ein beliebiger Datenumfang vorhanden ist. Anschließend widmen wir uns den Preiszeitreihen F1B und F1P.

3.7.1 Autoregressive Zeitreihen

Autoregressive Zeitreihen ohne Jumps

Wir beschäftigen uns zunächst mit autoregressiven Prozessen ohne gekoppelten Poissonprozess. Die Güte der angepassten verallgemeinerten Extremwertverteilung wird jeweils mit einem QQ-Plot visualisiert. Des Weiteren werden empirische und theoretische Verteilung gegeneinander geplottet und die Daten werden mit der entsprechenden Extremwertverteilungsannahme dem Kolmogorov-Smirnov-Test unterzogen. Für diese Untersuchungen betrachten wir die folgende Zeitreihe:

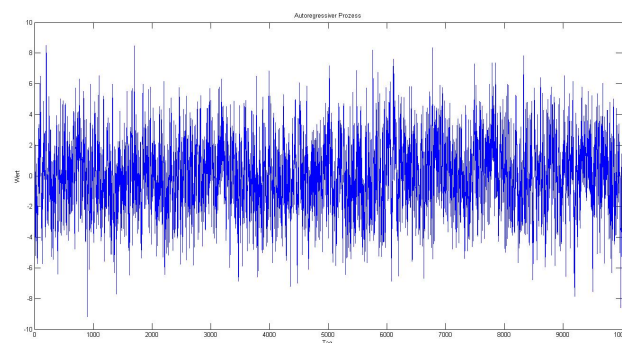


Abbildung 3.7: autoregressiver Prozess über 10.000 Tage

Durch die Erkenntnisse des vorherigen Kapitels, erwiesen sich eine Blocklänge von 7, 22 und 44 Tagen als sehr vielversprechend. Nach Anwendung des Verfahrens mit den jeweiligen Blocklängen erhalten wir folgende QQ-Plots:

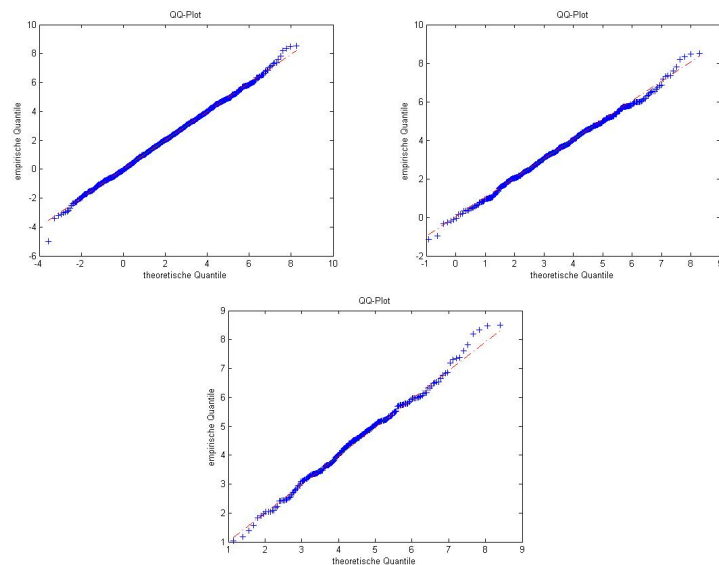


Abbildung 3.8: Vergleich der QQ-Plots mit einer Blocklänge von 7, 22 und 44 (von links)

In den Abbildungen ist zu erkennen, dass mit allen Blocklängen die theoretischen und die empirischen Quantile ab einem Wert von 7 immer mehr voneinander abweichen. Die Quantile mit einer Blocklänge von 7 scheinen auf den ersten Blick am wenigsten abzuweichen. Jedoch muss beachtet werden, dass in diesem Szenario deutlich mehr Werte für die Schätzung zur Verfügung standen. Zu den QQ-Plots ergaben sich die folgenden Verteilungen:

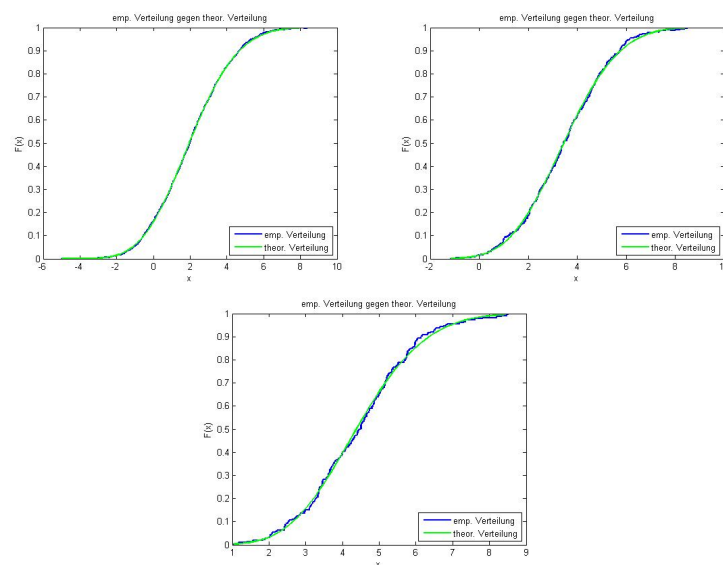


Abbildung 3.9: Vergleich der gefitteten Verteilung mit der empirischen Verteilung mit einer Blocklänge von 7, 22 und 44 (von links)

Während bei einer Blocklänge von 7 die empirische und die geschätzte Verteilung nahezu miteinander verschmelzen, kann bei den größeren Blocklängen kleinere Abweichungen erkannt werden, die mit Zunahme der Blocklänge immer deutlicher werden, was eine Folge der geringeren Datemenge ist. Dennoch ist bei allen drei Schätzungen die

Annäherung an die empirische Verteilung sichtbar zu erkennen.

Zum Abschluss dieser Untersuchung sind hier die folgenden Kenngrößen aufgelistet:

Blocklänge	Kolmogorov-Smirnov-Test	p -Level	$VaR_{99\%}$
7 Tage	0	0,9875	6,84
22 Tage	0	0,9652	7,56
44 Tage	0	0,9525	7,99

Mit einer Blocklänge von 7 Tagen erhalten wir das höchste p -Level und den kleinsten VaR_{99} . Mit Zunahme der Blocklänge wird das p -Level kleiner, allerdings erhöht sich der VaR_{99} , da dieser für einen größeren Block gilt.

Autoregressive Zeitreihen mit Jumps

Um die spontanen Preissprünge simulieren zu können, koppeln wir einen Poissonprozess an den autoregressiven Prozess. Wir betrachten nun die folgende Zeitreihe:

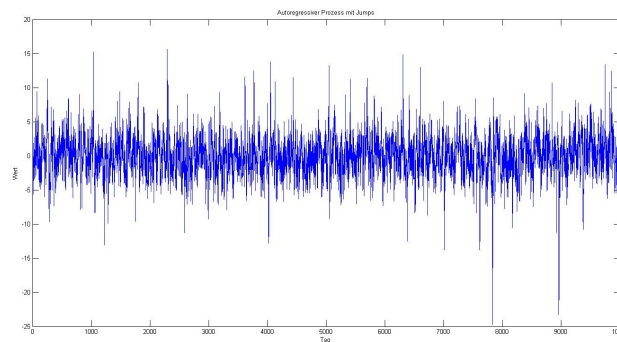


Abbildung 3.10: autoregressiver Prozess mit Jumps über 10.000 Tage

Für diese Zeitreihe zeigte sich die Wahl der Blocklängen 24, 40 und 55 als günstig. Letzterer ist sogar das globale Optimum über alle Blocklängen. Wie schon zuvor vergleichen wir, nach Anwendung des Verfahrens, die empirischen und die geschätzten Quantile miteinander. Es ergeben sich die folgenden QQ-Plots:

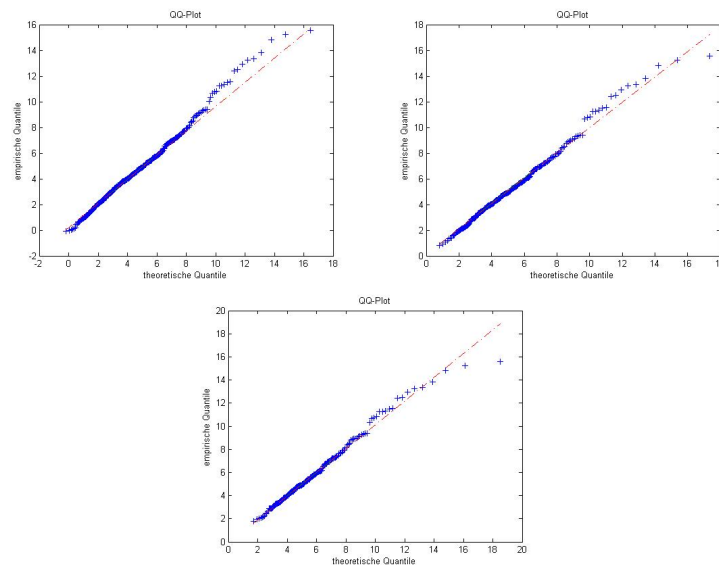


Abbildung 3.11: Vergleich der QQ-Plots mit einer Blocklänge von 24, 40 und 55 (von links)

Anders als bei autoregressiven Zeitreihen ohne Jumps weichen hier die Quantile bei kleinerer Blocklänge mehr voneinander ab, als bei größerer Blocklänge. Die Quantile der Blocklängen von 40 und 55 liegen ab einem Wert von 10 näher an der Winkelhalbierenden als bei einer Blocklänge von 24. Betrachten wir zudem noch die Verteilungen:

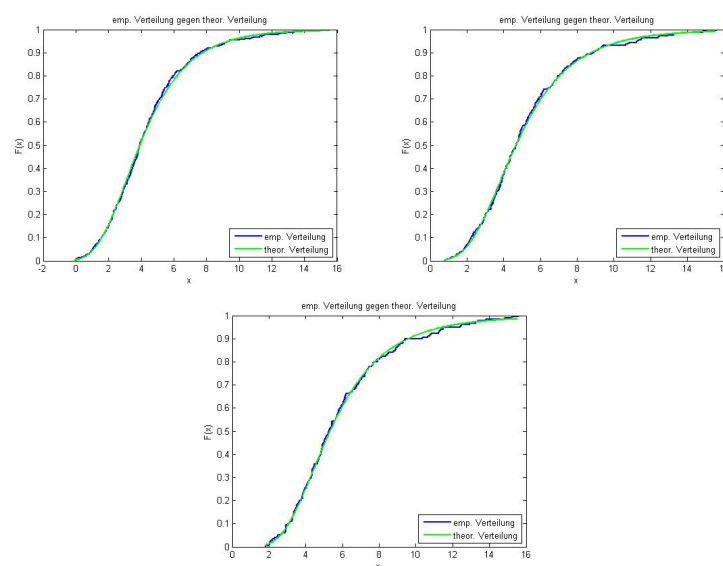


Abbildung 3.12: Vergleich der gefitteten Verteilung mit der empirischen Verteilung mit einer Blocklänge von 24, 40 und 55 (von links)

Ebenso wie bei autoregressiven Zeitreihen ohne Jumps nähert sich die geschätzte Extremwertverteilung sichtbar der empirischen Verteilung an. Bei den Blocklängen 40 und 55 können kleinere Abweichungen zwischen beiden Verteilungen erkannt werden, die jedoch auf den geringeren Datenumfang zurückzuführen sind.

Zum Schluss betrachten wir die Kenngrößen:

Blocklänge	Kolmogorov-Smirnov-Test	p -Level	$VaR_{99\%}$
24 Tage	0	0,9437	13,03
40 Tage	0	0,9898	14,72
55 Tage	0	0,9969	16,43

Wie schon mit autoregressiven Zeitreihen ohne Jumps erhalten wir mit einer höheren Blocklänge einen höheren VaR_{99} . Jedoch nimmt mit Zunahme der Blocklänge auch das p -Level zu und erreicht sein Maximum bei 55 Tagen, welches das globale Optimum über alle Blocklängen dieser Zeitreihe ist.

3.7.2 Preiszeitreihen

Phelix-Index Baseload

Nach der Untersuchung von stationären Prozessen kommen wir nun zu den praktischen Zeitreihen und beginnen mit dem Phelix-Index im Base:

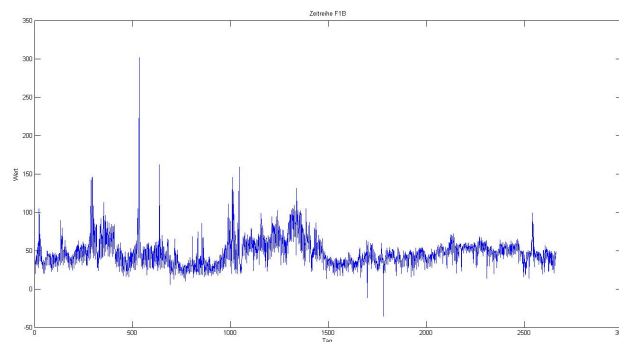


Abbildung 3.13: Phelix-Index im Base über 2662 Tage

Im vorherigen Kapitel haben wir die Blocklängen von 7 bzw. 24 Tage ausgewählt, die uns ein überzeugendes Ergebnis versprochen hatten. Betrachten wir zuerst die QQ-Plots:

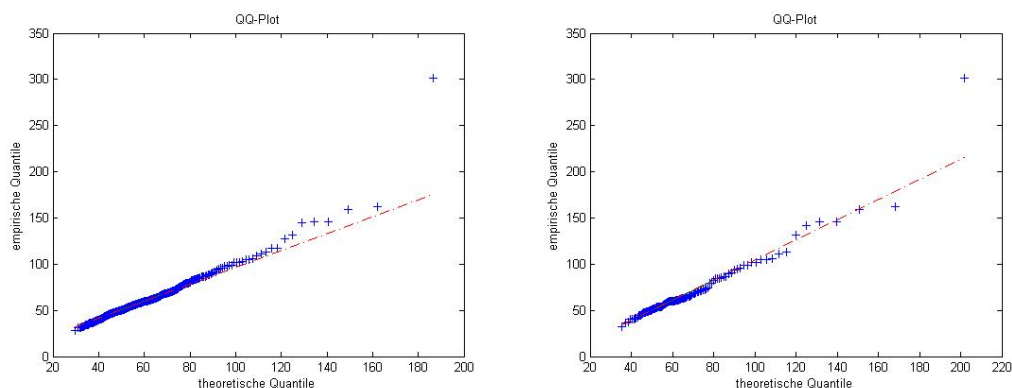


Abbildung 3.14: Vergleich der QQ-Plots mit einer Blocklänge von 7 und 24 Tagen (von links)

In der linken Abbildung (Blocklänge 7 Tage) weichen die Quantile ab einem Preis von ca. 120 EUR/MWh voneinander ab. Dagegen liegen die Quantile mit der höheren Blocklänge von 24 Tagen bis zu einem Preis von ca. 150 EUR/MWh zusammen. Das maximale Quantil von über 300 EUR/MWh kann jedoch mit beiden Blocklängen nicht abgeschätzt werden. Werfen wir nun einen Blick auf die geschätzten Verteilungen:

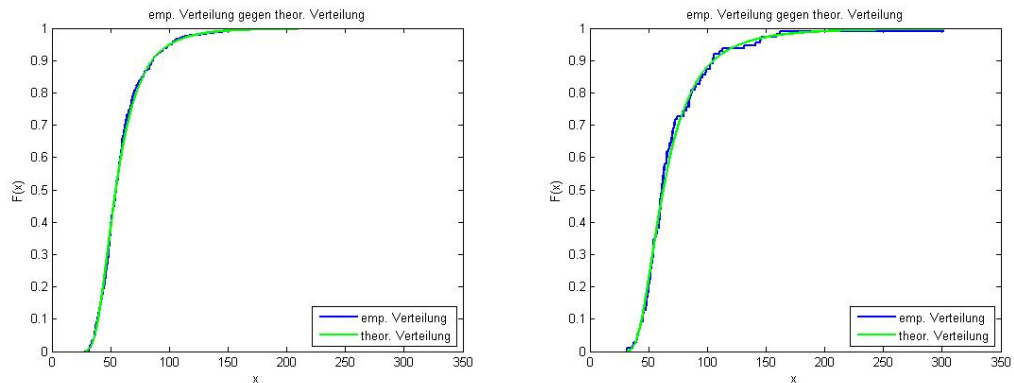


Abbildung 3.15: Vergleich der gefitteten Verteilung mit der empirischen Verteilung mit einer Blocklänge von 7 und 24 Tagen (von links)

Man erkennt deutlich, dass sich die geschätzten Verteilungen beider Blocklängen an die empirische Verteilung anpassen. Bei einem 24-Tage-Block sind kleine Abweichungen zu erkennen, die jedoch auf die geringere Datenmenge zurückzuführen sind. Abschließend untersuchen wir noch die Kenngrößen:

Blocklänge	Kolmogorov-Smirnov-Test	p -Level	$VaR_{99\%}$
7 Tage	0	0,6559	142,10 EUR/MWh
24 Tage	0	0,8010	196,32 EUR/MWh

Der 24-Tage-Block erreicht das höhere p -Level, da bei dem 7-Tage-Block auch viele kleine Datenpunkte in die Schätzung fließen, die keine Extremwerte sind.

Phelix-Index Peakload

Zum Abschluss dieses Kapitels wollen wir noch den Phelix-Index im Peak untersuchen:

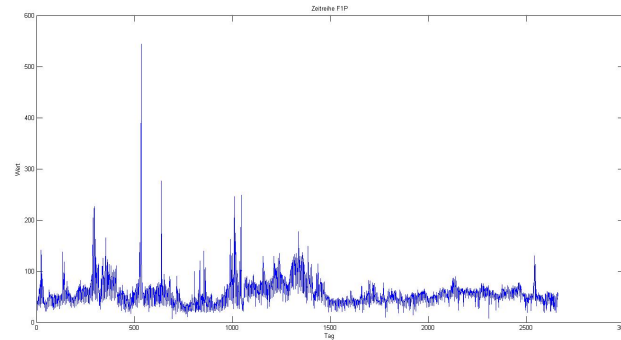


Abbildung 3.16: Phelix-Index im Peak über 2662 Tage

Ähnlich wie beim Base, ergeben sich hier die Blocklängen von 7 und 25 Tagen als lukrative Wahl. Dazu betrachten wir zunächst die QQ-Plots:

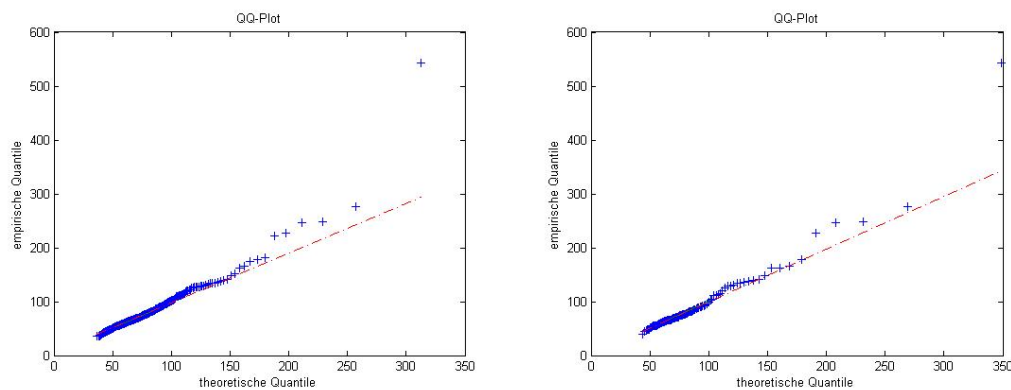


Abbildung 3.17: Vergleich der QQ-Plots mit einer Blocklänge von 7 und 25 Tagen (von links)

Bei der kleineren Blocklänge weichen die Quantile ab einem Preis von 170-180 EUR/MWh deutlich voneinander ab. Dieses Verhalten kann bei der größeren Blocklänge ebenfalls festgestellt werden, jedoch ist die Abweichung in der Preisspanne von 170-270 EUR/MWh kleiner als bei der kleineren Blocklänge. Das maximale Quantil von über 500 EUR/MWh kann jedoch mit keiner der beiden Blocklängen erreicht werden.

Sehen wir uns nun die Verteilungen an:

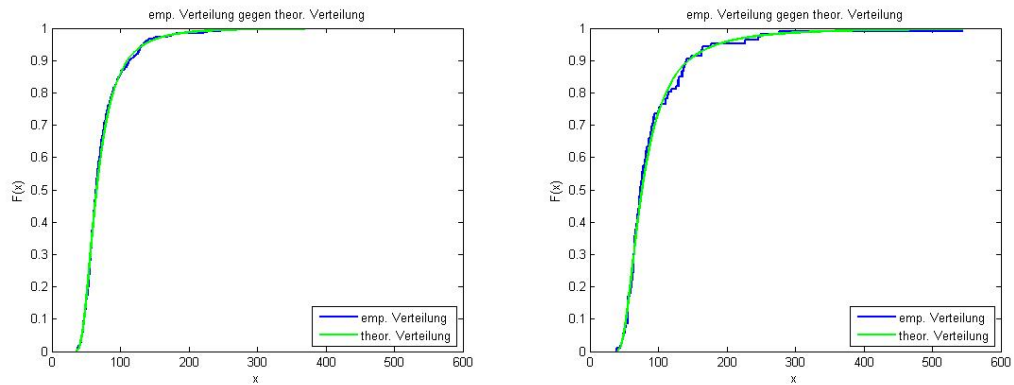


Abbildung 3.18: Vergleich der gefitteten Verteilung mit der empirischen Verteilung mit einer Blocklänge von 7 und 25 Tagen (von links)

Wie schon im Base kann auch hier eine deutliche Annäherung der geschätzten und empirischen Verteilung festgestellt werden. Bei dem 25er-Block sind wieder kleine Abweichungen zu erkennen, die jedoch wie schon beim Base auf einen geringeren Datenumfang zurückzuführen sind. Zum Abschluss werfen wir noch einen Blick auf die Kenngrößen:

Blocklänge	Kolmogorov-Smirnov-Test	p -Level	$VaR_{99\%}$
7 Tage	0	0,6746	213,76 EUR/MWh
25 Tage	0	0,9380	340,28 EUR/MWh

Ebenfalls wie beim Baseload hat auch hier die größere Blocklänge das höhere p -Level, was durch die höheren Blockmaxima begründet wird. Mit Zunahme der Blocklänge steigt selbstverständlich auch der VaR_{99} , da dieser auch für einen längeren Zeitraum gilt.

4 Peak over Threshold (POT)

In diesem Kapitel wenden wir uns einer alternativen Methode der Extremwerttheorie zu. Die Theorie der Schwellenwertüberschreitung basiert auf der Verteilung der Exzesse, der verallgemeinerten Pareto-Verteilung [6].

4.1 Verallgemeinerte Pareto-Verteilungen

Ausgangspunkt sind unsere ursprünglichen Daten X_1, \dots, X_n , die einer unbekannten Verteilung F unterliegen. Durch Festlegung eines geeigneten Schwellenwertes $u \in \mathbb{R}$ interessieren wir uns nur für die Verteilung der Exzesse $Y = X - u$ unter der Bedingung, dass $X > u$. Wir betrachten somit nur den rechten Teil ab einer Schranke u der Verteilung F . Gesucht ist nun die Verteilung $F_u(x)$, die uns die Wahrscheinlichkeit des folgenden Ereignisses berechnet:

$$F_u(x) = P(X - u \leq x | X > u). \quad (4.1)$$

Laut Satz der totalen Wahrscheinlichkeit gilt für zwei Ereignisse A und B :

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \quad (4.2)$$

Für (4.1) ergibt sich somit:

$$F_u(x) = \frac{P(X \leq u + x, X > u)}{P(X > u)} = \frac{F(u + x) - F(u)}{1 - F(u)}, \quad (4.3)$$

für $u < x^*$ und $x \geq 0$ [14]. Auf diese Darstellungsform werden wir im späteren Verlauf des Kapitels nochmals zurückkommen. Aus (4.3) kann noch nicht auf die Gestalt von $F_u(x)$ geschlossen werden, da die Verteilung F im Allgemeinen unbekannt ist. Eine wichtige Verteilung der Peak over Threshold ist die verallgemeinerte Pareto-Verteilung.

Definition 4.1 Die Verteilungsfunktion der **Verallgemeinerten Pareto-Verteilung** wird definiert durch

$$P_{\gamma, \beta}(x) = \begin{cases} 1 - \left(1 + \gamma \frac{x}{\beta}\right)^{-\frac{1}{\gamma}}, & \gamma \neq 0 \\ 1 - e^{-\frac{x}{\beta}}, & \gamma = 0 \end{cases}, \quad (4.4)$$

wobei $\beta > 0$, und $x > 0$ wenn $\gamma \geq 0$ und $0 \leq x \leq -\frac{\beta}{\gamma}$ wenn $\gamma < 0$ [6].

Der Fall für $\gamma = 0$ muss wieder, analog zur verallgemeinerten Extremwertverteilung, über den Grenzwert betrachtet werden.

$$P_{0,\beta}(x) = \lim_{\gamma \rightarrow 0} \left(1 - \left(1 + \gamma \frac{x}{\beta} \right)^{-\frac{1}{\gamma}} \right) = 1 - e^{-\frac{x}{\beta}}, \quad t > 0.$$

$P_{0,\beta}$ stimmt somit mit der Exponentialverteilung mit Parameter $\frac{1}{\beta}$ überein. Für den Fall, dass $\gamma < 0$ nennt man $P_{\gamma,\beta}$ auch Pareto-Typ-II-Verteilung, für $\gamma > 0$ allgemeine Pareto-Verteilung.

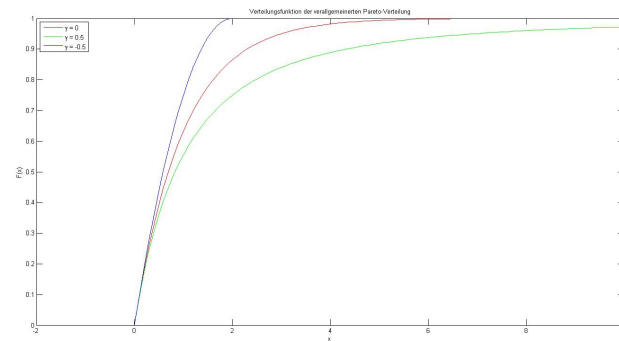


Abbildung 4.1: Graphen der Pareto-Verteilungen mit $\beta = 1$

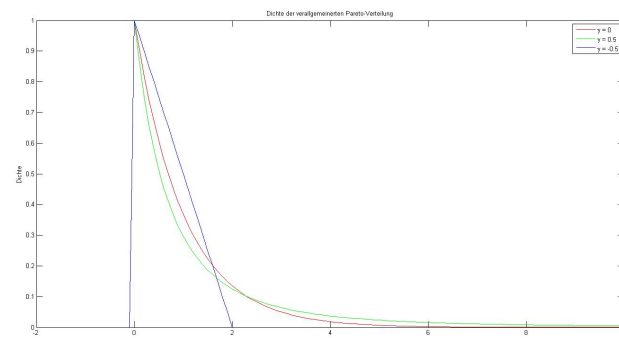


Abbildung 4.2: Graphen der Pareto-Dichten mit $\beta = 1$

Nun ist die Frage zu klären, ob die Verteilung der Exzesse $X - u$, unter der Bedingung, dass $X > u$, durch die verallgemeinerte Pareto-Verteilung ausgedrückt werden kann. Antwort auf diese Frage gibt uns das Theorem von Pickands-Balkeema-de Haan.

4.2 Satz von Pickands-Balkema-de Haan

Satz 4.2 (Satz von Pickands-Balkeema-de Haan) *Sei X eine Zufallsgröße mit Verteilungsfunktion F , die im rechten Endpunkt x^* stetig ist. Dann liegt F im Max-Anziehungsbereich von G_γ genau dann, wenn es eine positive messbare Funktion $\beta(u)$ gibt mit*

$$\lim_{u \rightarrow x^*} \sup_{t \in [0, x^* - u]} |P(X - u \leq t | X > u) - P_{\gamma, \beta(u)}(x)| = 0. \quad (4.5)$$

Das heißt, es gilt die Approximation $P(X - u \leq x | X > u) \approx P_{\gamma, \beta(u)}(x)$, falls F im Max-Anziehungsbereich von G_γ liegt und wir können die verallgemeinerte Pareto-Verteilung als Näherung für die Verteilung der Exzesse benutzen [5]. Für diesen wichtigen Satz werden wir im Folgenden eine Beweisskizze angeben, die auf dem Satz von Fisher-Tippett beruht [6]:

Beweisskizze Satz von Pickands-Balkeema-de Haan: Wenn F die Verteilungsfunktion von X ist, dann gilt für $n \rightarrow \infty$ nach Satz von Fisher-Tippett

$$F \in D(G_\gamma) \Leftrightarrow F^n(x) \approx \exp \left(- \left(1 + \gamma \left(\frac{x - \mu}{\beta} \right) \right)^{-\frac{1}{\gamma}} \right),$$

für $\mu, \sigma > 0$ und $\gamma \in \mathbb{R}$. Nun logarithmieren wir beide Seiten und erhalten:

$$n \log F(x) \approx - \left(1 + \gamma \left(\frac{x - \mu}{\beta} \right) \right)^{-\frac{1}{\gamma}}. \quad (4.6)$$

Man kann über eine Taylorreihenentwicklung zeigen, dass für große Werte x gilt

$$\log F(x) \approx -(1 - F(x)). \quad (4.7)$$

Nun wird (4.7) in (4.6) eingesetzt und nach $F(x)$ umgestellt. Dabei erhalten wir

$$F(x) \approx 1 - \frac{1}{n} \left(1 + \gamma \left(\frac{x - \mu}{\beta} \right) \right)^{-\frac{1}{\gamma}}.$$

Somit ergibt sich für großes u :

$$F(u) \approx 1 - \frac{1}{n} \left(1 + \gamma \left(\frac{u - \mu}{\beta} \right) \right)^{-\frac{1}{\gamma}} \quad (4.8)$$

und für $x > 0$:

$$F(u+x) \approx 1 - \frac{1}{n} \left(1 + \gamma \left(\frac{u+x-\mu}{\beta} \right) \right)^{-\frac{1}{\gamma}}. \quad (4.9)$$

Nun benötigen wir die Exzessverteilung. Dabei betrachte man die Form (4.3), in die (4.8) und (4.9) eingesetzt wird:

$$\begin{aligned} F_u(x) &= \frac{F(u+x) - F(u)}{1 - F(u)} \\ &\approx \frac{1 - \frac{1}{n} \left(1 + \gamma \left(\frac{u+x-\mu}{\beta} \right) \right)^{-\frac{1}{\gamma}} - \left(1 - \frac{1}{n} \left(1 + \gamma \left(\frac{u-\mu}{\beta} \right) \right)^{-\frac{1}{\gamma}} \right)}{1 - \left(1 - \frac{1}{n} \left(1 + \gamma \left(\frac{u-\mu}{\beta} \right) \right)^{-\frac{1}{\gamma}} \right)}. \end{aligned}$$

Durch einfaches Zusammenfassen und Vereinfachen erhalten wir das folgende Resultat:

$$F_u(x) = 1 - \left(1 + \frac{\gamma x}{\beta + \gamma(u - \mu)} \right)^{-\frac{1}{\gamma}}, \quad (4.10)$$

was der Verallgemeinerten Pareto-Verteilung $P_{\gamma, \beta + \gamma(u - \mu)}(x)$ entspricht. Daraus folgt nun:

$$F_u(x) = P_{\gamma, \beta + \gamma(u - \mu)}(x)$$

mit

$$\beta(u) = \beta + \gamma(u - \mu).$$

□

Erinnern wir uns nochmals an die Verteilung der Blockmaxima im Abschnitt 3.3. Es besteht ein Zusammenhang zwischen dieser Verteilung und der Verallgemeinerten Pareto-Verteilung, den wir nun herleiten werden. Man betrachte zunächst die Verteilungsfunktion der Blockmaxima

$$G_\gamma(t) = \exp \left\{ -(1 + \gamma t)^{\frac{-1}{\gamma}} \right\}$$

und bilde auf beiden Seiten den Logarithmus:

$$\log G_\gamma(t) = -(1 + \gamma t)^{\frac{-1}{\gamma}}.$$

Addiert man nun auf beiden Seiten 1, erhält man:

$$1 + \log G_\gamma(t) = 1 - (1 + \gamma t)^{\frac{-1}{\gamma}} = P_\gamma(t),$$

mit $t = \frac{x}{\beta}$.

Es besteht also der folgende Zusammenhang zwischen der verallgemeinerten Extremwertverteilung und der verallgemeinerten Pareto-Verteilung [6]:

$$P_\gamma(t) = 1 + \log G_\gamma(t). \quad (4.11)$$

Durch (4.11) gilt, dass die Verteilung der Maxima von unabhängig, identisch verteilten Zufallsgrößen gleich $G_\gamma(t)$ ist, genau dann, wenn die Exzesse bzgl. einer Schranke u derselben Zufallsgrößen die Verteilungsfunktion $P_\gamma(t) = 1 + \log G_\gamma(t)$ haben. Da im Abschnitt 3.4 gezeigt werden konnte, dass das Extremwertverhalten von stationären Zeitreihen und von unabhängig, identisch verteilten Zufallsgrößen übereinstimmt und beide Verteilungen der jeweiligen Maxima gegen dieselbe Extremwertverteilung konvergieren, können wir dieses Ergebnis auch auf das Verfahren der Schwellenwertüberschreitung übertragen. Das heißt, Peak over Threshold kann ebenfalls für stationäre Zeitreihen angewendet werden, die die Eigenschaften aus 3.4 erfüllen.

4.3 Wahl des Schwellenwertes

Bei Peak over Threshold ist die Wahl des Schwellenwertes u von großer Bedeutung. Die gewählte Schranke bestimmt sozusagen die Anzahl der Daten, die für die Anpassung an die Verallgemeinerte Pareto-Verteilung berücksichtigt werden. Das Problem bei der

Festlegung von u ist einen geeigneten Kompromiss zu finden. Wählt man die Schwelle sehr groß, so werden auch nur wirklich große Überschreitungen berücksichtigt. Jedoch bleiben am Ende nur wenige Daten übrig, die für die Anpassung an eine Verteilung herangezogen werden können. Dadurch unterliegen die ausgewählten Daten unter Umständen einer hohen Varianz. Auf der anderen Seite ist eine kleine Schranke ebenfalls von Nachteil. Für die Anpassung an die Verteilung liegen dann zugegeben viele Daten vor, jedoch muss man beachten, dass der Großteil der Daten, keine extremen Ereignisse darstellen, wodurch die Aussage des Pickands-Balkeema-de Haan-Theorem nicht mehr greift. Gefordert wird dort, dass die Schranke u nahe dem rechten Endpunkt x^* liegen soll.

Eine Idee zur Wahl eines geeigneten Schwellenwert sind die empirischen Quantile der Ausgangsdaten. Wir ordnen zunächst unsere Realisierungen der Zeitreihe der Größe nach:

$$X_{(1)} \leq \dots \leq X_{(n)}.$$

Anschließend multipliziert man die Länge der Zeitreihe n mit einer Zahl $0 < p < 1$. Das Produkt wird aufgerundet zu $[np]$ und ergibt das empirische p -Quantil der Ausgangsdaten bzw. den Datenpunkt $X_{i_{[np]}}$, der mit einer Wahrscheinlichkeit von p nicht überschritten wird. Für Peak over Threshold ist es nötig, dass p nahe 1 ist, damit das Pickands-Balkeema-de Haan-Theorem greift. In unserem Fall sollte der Schwellenwert zwischen dem 0,9- und 0,99-Quantil der Daten liegen.

Diese Idee hat den Vorteil, dass sie uns durch einfache Mittel einen geeigneten Schwellenwert gibt, der extreme Ereignisse von alltäglichen trennt und dennoch einen gewissen Datenumfang bewahrt. Selbst bei der Wahl des 0,99-Quantils erhalten wir bei unseren Zeitreihen F1B und F1P, die eine Länge von $n = 2662$ Börsentagen haben, immer noch 26 Datenpunkte, die für die Anpassung an die Verallgemeinerte Pareto-Verteilung herangezogen werden können. Bei der Wahl des 0,98-Quantil sind es 53 Datenpunkte, mit denen wir statistisch glaubwürdige Aussagen treffen können.

In der Literatur wird eine weitere Möglichkeit vorgestellt, die eine Hilfestellung für dieses Problem bietet. Man nennt diese Methode Mean Excess Plot [5] oder auch Mean Residual Life Plot. Dazu benötigen wir die mittlere Exzess-Funktion:

$$e(u) = E(X - u | X > u).$$

Man kann zeigen, dass diese, wenn das Theorem von Pickands-Balkeema-de Haan greift, einen linearen Verlauf hat [5]. Somit sollte dann auch die empirische mittlere Exzessfunktion $e_n(u)$ näherungsweise linear verlaufen, wobei

$$e_n(u) = \frac{\sum_{i=1}^n (X_i - u) I_{\{X_i > u\}}}{\sum_{i=1}^n I_{\{X_i > u\}}},$$

mit $I_{\{X_i > u\}}$ als Indikatorfunktion [5].

Allerdings ist die Identifikation der Stelle u , ab dem $e_n(u)$ näherungsweise linear verlaufen soll, in der Praxis oft nicht eindeutig. Für unsere Preiszeitreihen F1B und F1P ergeben sich folgende Exzessplots:

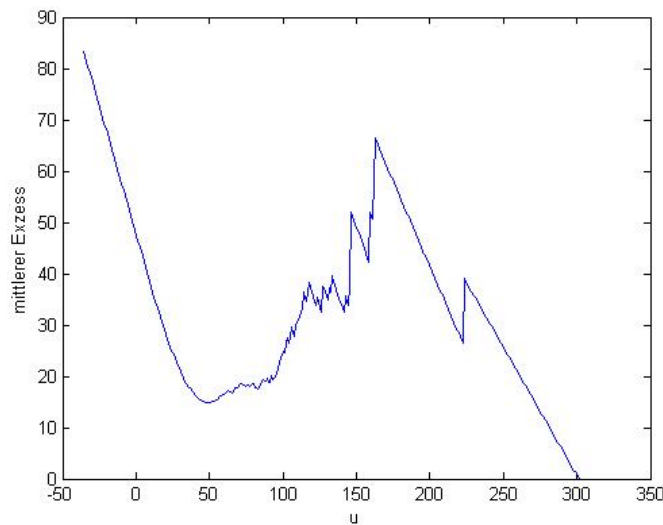


Abbildung 4.3: Mean Excess Plot für F1B

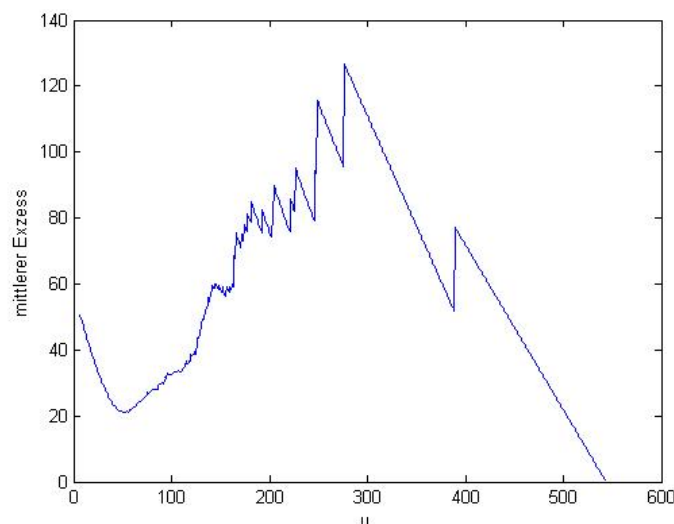


Abbildung 4.4: Mean Excess Plot für F1P

Sowohl in Abbildung 4.3 als auch in Abbildung 4.4 ist es nicht möglich einen eindeutigen Schwellenwert zu finden, weswegen dieses Verfahren für unser Problem ungeeignet

ist.

4.4 Schätzung der Parameter und des Value at Risk

Bevor wir Peak over Threshold auf praxisrelevante Daten anwenden können, muss noch die Schätzung der Parameter betrachtet werden, die in diesem Abschnitt genauer beschrieben werden soll.

Ausgangspunkt sind unsere Daten X_1, \dots, X_n , von denen wir jedoch nur die Datenpunkte X_{i_1}, \dots, X_{i_k} berücksichtigen, die über einer vorher gewählten Schwelle u liegen. Wir definieren die Exzesse als

$$Y_1 = X_{i_1} - u, \dots, Y_k = X_{i_k} - u. \quad (4.12)$$

Gesucht ist nun der Parametervektor $\Pi = (\gamma, \beta) \in \mathbb{R}^2$, so dass die Gestalt der verallgemeinerten Pareto-Verteilung für deren Realisierungen Y_1, \dots, Y_k beschrieben werden kann. Für diese Aufgabe bieten sich gleich mehrere Schätzmethoden an, die wir im Folgenden untersuchen wollen.

4.4.1 Maximum-Likelihood-Methode

Ein klassisches Schätzverfahren in der Statistik für unbekannte Parameter ist die Maximum-Likelihood-Methode, die wir schon im Abschnitt 3.5 für die Blockmaxima untersucht haben. Für die Anwendung wird zunächst die Dichte $f_\Pi(y)$ der verallgemeinerten Pareto-Verteilung benötigt. Diese erhalten wir, indem von $P_\Pi(y)$ die erste Ableitung gebildet wird. Für $\gamma \neq 0$ ergibt sich somit:

$$f_\Pi(y) = \begin{cases} \frac{1}{\beta} \left(1 + \frac{\gamma y}{\beta}\right)^{-\frac{1}{\gamma}-1}, & y \geq 0 \\ 0, & \text{sonst.} \end{cases}$$

Für den Fall, dass $\gamma = 0$, muss wieder der Grenzwert betrachtet werden:

$$\begin{aligned}
f_{0,\beta}(y) &= \lim_{\gamma \rightarrow 0} \frac{1}{\beta} \left(1 + \frac{\gamma y}{\beta}\right)^{-\frac{1}{\gamma}-1} \\
&= \lim_{\gamma \rightarrow \infty} \frac{1}{\beta} \left(1 - \frac{-y}{\gamma\beta}\right)^{-\gamma-1} \\
&= \frac{1}{\beta} e^{-\frac{y}{\beta}}.
\end{aligned}$$

Mit Hilfe der Dichten kann nun die Log-Likelihoodfunktion gebildet werden [7]:

$$\log l(\gamma, \beta) = -k \log \beta - \left(1 + \frac{1}{\gamma}\right) \sum_{i=1}^k \log \left(1 + \frac{\gamma y_i}{\beta}\right), \quad 1 + \frac{\gamma y_i}{\beta} > 0 \quad \forall i.$$

Für $\gamma = 0$ ergibt sich [7]:

$$\log l(0, \beta) = -k \log \beta - \frac{1}{\beta} \sum_{i=1}^k y_i.$$

Um nun den Maximum-Likelihood-Schätzer berechnen zu können, muss der Gradient von $l(\Pi)$ gleich 0 gesetzt werden:

$$\nabla l = \begin{pmatrix} \frac{\partial l}{\partial \gamma} \\ \frac{\partial l}{\partial \beta} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (4.13)$$

Das entstandene nichtlineare Gleichungssystem (4.13) kann, wie schon im Abschnitt 3.5, im Allgemeinen nicht exakt gelöst werden. Es muss an dieser Stelle wieder auf numerische Lösungsverfahren zurückgegriffen werden. Die Lösung von (4.13) ist unser gesuchter Parametervektor:

$$\hat{\Pi} = (\hat{\gamma}, \hat{\beta}) = \operatorname{argmax} l(\gamma, \beta).$$

Mit den gewonnenen Parametern kann nun der Value-at-Risk geschätzt werden. Wir wählen uns eine Wahrscheinlichkeit α , mit der der Wert $VaR_{1-\alpha}$ überschritten wird. Zur Berechnung von $VaR_{1-\alpha}$ benutzen wir das Resultat des Satzes von Pickands-Balkeema-de Haan [5]:

$$P(X - u > x | X > u) \approx \left(1 + \frac{\gamma x}{\beta}\right)^{-\frac{1}{\gamma}}.$$

Wir setzen $x = VaR_{1-\alpha} - u$ und es folgt aus dem Satz der totalen Wahrscheinlichkeit (4.2):

$$P(X > VaR_{1-\alpha}) \approx P(X > u) \left(1 + \gamma \frac{VaR_{1-\alpha} - u}{\beta}\right)^{-\frac{1}{\gamma}}.$$

$P(X > VaR_{1-\alpha})$ ist laut Voraussetzung α . Wenn man nun die obige Gleichung nach $VaR_{1-\alpha}$ umstellt, erhält man [5]:

$$Va\hat{R}_{1-\alpha} = u + \frac{\beta}{\gamma} \left(\left(\frac{P(X > u)}{\alpha} \right)^{\gamma} - 1 \right), \quad (4.14)$$

wobei $P(X > u) = \frac{1}{n} \sum_{i=1}^n 1_{X_i > u}$ gesetzt werden kann. Der hergeleitete $VaR_{1-\alpha}$ gilt für einen Tag, wenn man Tagespreise als Daten hat. Im Vergleich zum $VaR_{1-\alpha}$ der Blockmaxima ist der Value at Risk von Peak over Threshold deutlich niedriger, da dieser nur für einen Tag gilt.

Unabhängig von der Value-at-Risk-Schätzung können noch zwei andere Schätzverfahren für die Parameter γ und β verwendet werden.

4.4.2 Momentenmethode

Als erstes alternatives Verfahren zum Schätzen des Parametervektors $\Pi = (\gamma, \beta)$ betrachten wir die Momentenmethode, bei der die Momente der theoretischen Verteilung durch die gesuchten Parameter ausgedrückt werden können. Wenn man die Momente anschließend durch ihre empirischen Schätzer ersetzt, kann nach den gesuchten Parametern umgestellt werden.

Zunächst benötigen wir den folgenden Satz [8]:

Satz 4.3 Sei Y gemäß einer verallgemeinerten Pareto-Verteilung verteilt mit Parameter γ und β und sei $r \in \mathbb{N}$. Für $\gamma < \frac{1}{r}$ gilt dann:

$$E(Y^r) = \frac{\beta^r}{\gamma^{r+1}} \frac{\Gamma(\gamma^{-1} - 1)}{\Gamma(\gamma^{-1} + 1)} r!.$$

Mit Hilfe dieses Satzes kann gezeigt werden, dass sich das erste und zweite Moment der verallgemeinerten Pareto-Verteilung wie folgt darstellen lässt [7]:

$$E(Y) = \frac{\beta}{1-\gamma}$$

$$Var(Y) = \frac{\beta^2}{(1-\gamma)^2(1-2\gamma)}.$$

Nun müssen die theoretischen Momente durch ihre empirischen Schätzer ersetzt werden. Man benutzt dabei $\bar{y} = \frac{1}{k} \sum_{i=1}^k y_i$ und $S^2 = \frac{1}{k-1} \sum_{i=1}^k (y_i - \bar{x})^2$. Nach dem Einsetzen der empirischen Schätzer und anschließendem Umstellen nach γ und β erhält man folgende Momentenschätzer [7]:

$$\hat{\gamma} = \frac{1}{2} \left(1 + \frac{\bar{y}^2}{S^2} \right)$$

$$\hat{\beta} = \frac{\bar{y}}{2} \left(1 - \frac{\bar{y}^2}{S^2} \right).$$

4.4.3 Wahrscheinlichkeitsgewichtete Momentenmethode (PWM)

Die im letzten Unterabschnitt besprochene Momentenmethode kann erweitert werden, indem man die theoretischen Momente mit Wahrscheinlichkeiten wichtet. Somit haben extrem große Werte eine geringe Auswirkung auf die Schätzung der Parameter, sodass die PWM, gerade für kleine Stichproben, zuverlässigere Schätzungen gibt (Hosking und Wallis 1987 [17]). Für eine Zufallsgröße Y mit Verteilungsfunktion $F(Y)$ sind die wahrscheinlichkeitsgewichteten Momente definiert als [7]:

$$M_{p,r,s} = E(Y^p F(Y)^r (1 - F(Y))^s), \quad (4.15)$$

wobei $p, r, s \in \mathbb{R}$. Mit $M_{p,0,0}$ erhalten wir den bekannten Momentenschätzer aus dem vorherigen Unterabschnitt. Nach [7] ist es günstig für die Schätzung der Parameter γ und β $p = 1$ und $r = 0$ zu wählen. Man kann mit (4.15) zeigen, dass

$$\omega_s = M_{1,0,s} = E(Y(1 - F(Y))^s) = \frac{\beta}{(s+1)(s+1-\gamma)}, \quad \gamma < 1. \quad (4.16)$$

Beweis: Nachzulesen in [9].

□

Für $s = 0$ und $s = 1$ erhält man schließlich:

$$\omega_0 = \frac{\beta}{1 - \gamma}$$

$$\omega_1 = \frac{\beta}{2(2 - \gamma)}.$$

Durch Umstellen nach γ und β erhalten wir:

$$\gamma = 2 - \frac{\omega_0}{\omega_0 - 2\omega_1}$$

$$\beta = \frac{2\omega_0\omega_1}{\omega_0 - 2\omega_1}.$$

Wenn man nun für ω_0 und ω_1 die jeweiligen empirischen Schätzer einsetzt, erhält man die gesuchten Schätzer $\hat{\gamma}$ und $\hat{\beta}$. In [9] stellt sich heraus, dass die unverzerrten Schätzer für ω_0 und ω_1 wie folgt aussehen:

$$\omega_0 = \frac{1}{k} \sum_{i=1}^k y_i$$

$$\omega_1 = \frac{1}{k} \sum_{i=1}^k \frac{k-i}{k-1} y_i,$$

wobei man ω_0 als Exzessmittelwert über der Schwelle u interpretieren kann.

4.5 Anwendung des Verfahrens

In diesem Abschnitt wollen wir nun die Theorie der Schwellenwertüberschreitung auf Zeitreihen anwenden. Zunächst werden autoregressive Zeitreihen betrachtet, um zu sehen, wie sich das Verfahren verhält, wenn die Stationarität, die Clusterfreiheit, eine schwache Abhängigkeitsstruktur und zudem ein beliebiger Datenumfang vorhanden ist, bevor wir zu unseren bekannten Preiszeitreihen kommen.

4.5.1 Autoregressive Zeitreihen

Autoregressive Zeitreihen ohne Sprünge

Wir betrachten zunächst autoregressive Prozesse ohne gekoppelten Poissonprozess. Die Güte der angepassten verallgemeinerten Pareto-Verteilung wird jeweils mit einem QQ-Plot visualisiert. Des Weiteren werden empirische und theoretische Verteilung gegeneinander geplottet und die Daten werden mit verallgemeinerter Pareto-Verteilungsannahme dem Kolmogorov-Smirnov-Test unterzogen.

Wir simulieren nun einen autoregressiven Prozess gemäß der Bildungsvorschrift $X_t = \alpha X_{t-1} + \varepsilon_t$, wobei $\alpha = 0,8$ und $\varepsilon_t \sim N(0,2)$:

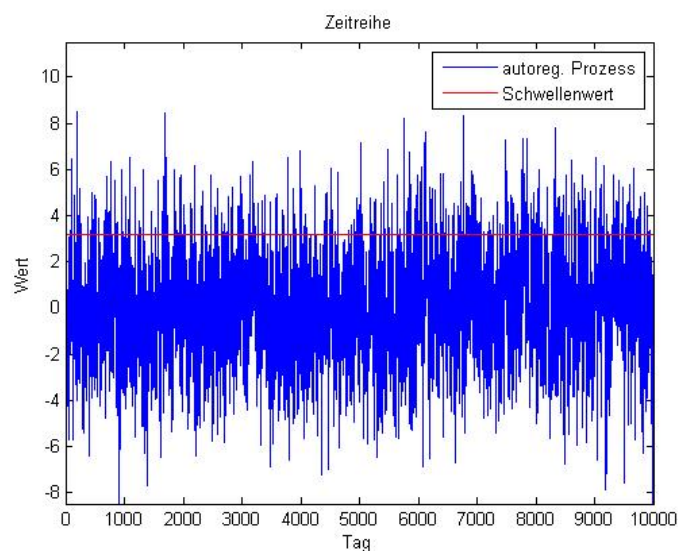


Abbildung 4.5: autoregressiver Prozess über 10.000 Tage mit 0,91-Quantil der Daten als Schwellenwert

Der Schwellenwert u ist hier das 0,91-Quantil der Daten, wodurch alle Werte in Betracht kommen, die größer als u sind. Die Schätzung der Parameter, in die nun die Exzesse einfließen, wird mit der Maximum-Likelihood-, der Momentenmethode, sowie den wahrscheinlichkeitsgewichteten Momenten vollzogen. Vergleichen wir anschließend die empirischen Quantile der Daten mit den theoretischen Quantilen der geschätzten Verteilung, so ergibt sich:

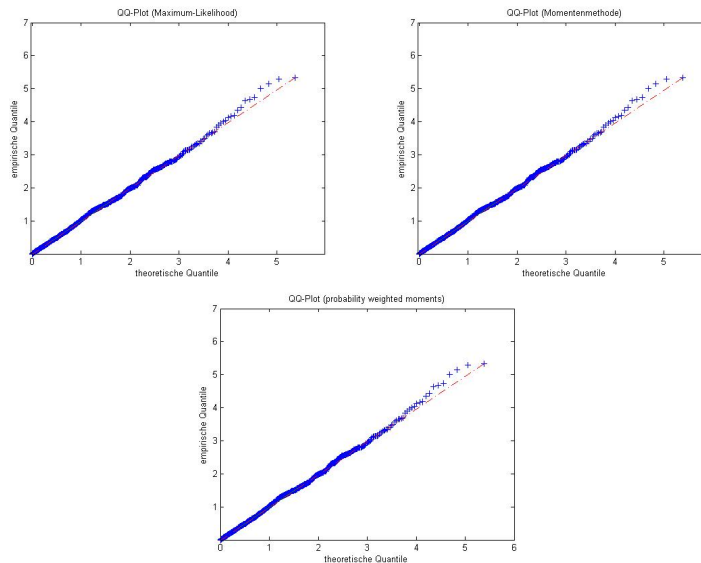


Abbildung 4.6: Vergleich der QQ-Plots mit Maximum-Likelihood, Momentenmethode und wahrscheinlichkeitsgewichteten Momenten

Man kann erkennen, dass sowohl die empirischen, als auch die theoretischen Quantile bis zu einer Überschreitung von ca. 4,5 bei allen Schätzmethodiken in etwa übereinstimmen. Bei Überschreitungen größer als 4,5 weichen die Quantile leicht voneinander ab, fallen dann aber zum Schluss wieder zusammen. Dass das größte Quantil in solch einem Fall wieder auf die Winkelhalbierende fällt ist statistisch gesehen ein Zufallstreffer und trägt nicht zur Güte des zu Grunde liegenden Modells bei. Interessant ist auch der Vergleich von empirischer und theoretischer Verteilung:

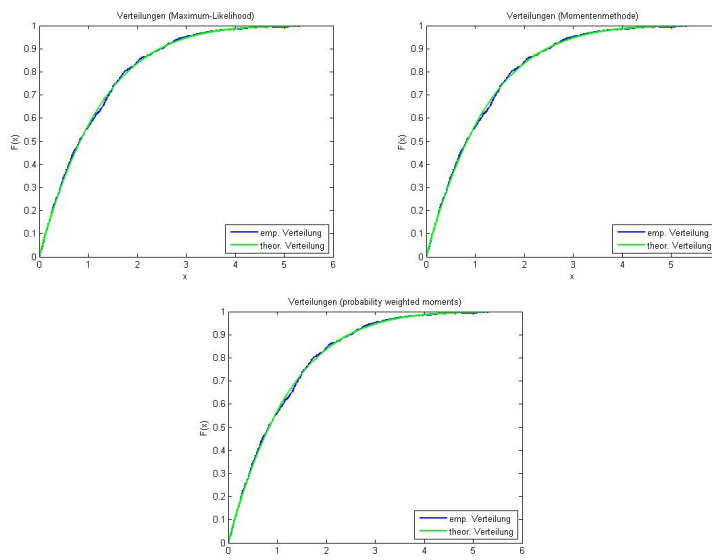


Abbildung 4.7: Vergleich der Verteilungen mit Maximum-Likelihood, Momentenmethode und wahrscheinlichkeitsgewichteten Momenten

Auch in diesem Vergleich ist deutlich zu erkennen, dass die theoretische Verteilung sich der empirischen Verteilung annähert. Zwischenzeitlich sind minimale Unterschiede bei

allen Schätzverfahren zuerkennen, die aber praktisch irrelevant sind. Diese minimalen Unterschiede sehen wir auch bei den Kenngrößen:

Schätzverfahren	Kolmogorov-Smirnov-Test	p -Level	$VaR_{99\%}$
Maximum-Likelihood	0	0,6613	5,54
Momentenmethode	0	0,6527	5,54
PWM	0	0,6531	5,54

Das p -Level und der $VaR_{99\%}$ sind bei allen drei Methoden ähnlich.

Erhöhen wir nun den Schwellenwert u auf das 0,99-Quantil der Daten und untersuchen alle drei Schätzmethoden:

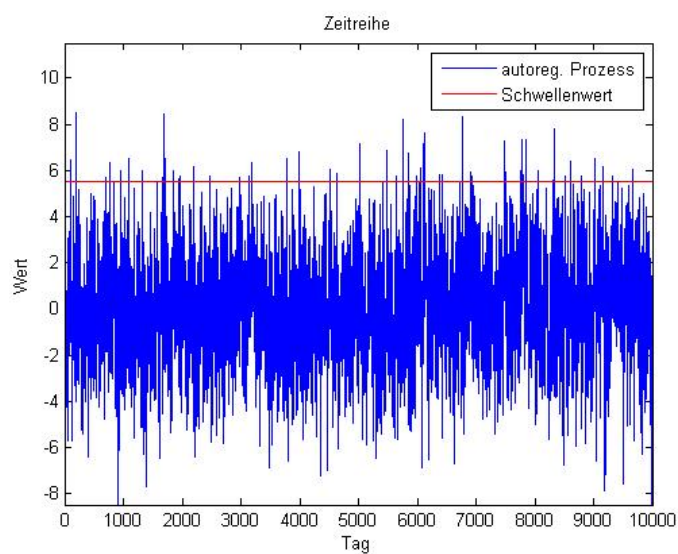


Abbildung 4.8: autoregressiver Prozess über 10.000 Tage mit 0,99-Quantil der Daten als Schwellenwert

Untersuchen wir zunächst die Quantile:

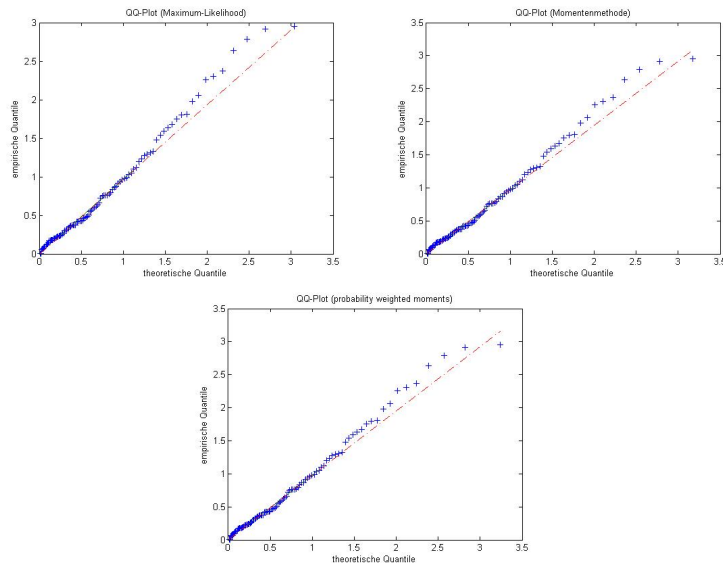


Abbildung 4.9: Vergleich der QQ-Plots mit Maximum-Likelihood, Momentenmethode und wahrscheinlichkeitsgewichteten Momenten

Bei allen drei Schätzverfahren ist zu beobachten, dass die Quantile ab einem Wert von ca. 1,5 von der Winkelhalbierenden abweichen, jedoch kann bei Maximum-Likelihood eine etwas größere Abweichung beobachtet werden.

Vergleichen wir nun die Verteilungen:

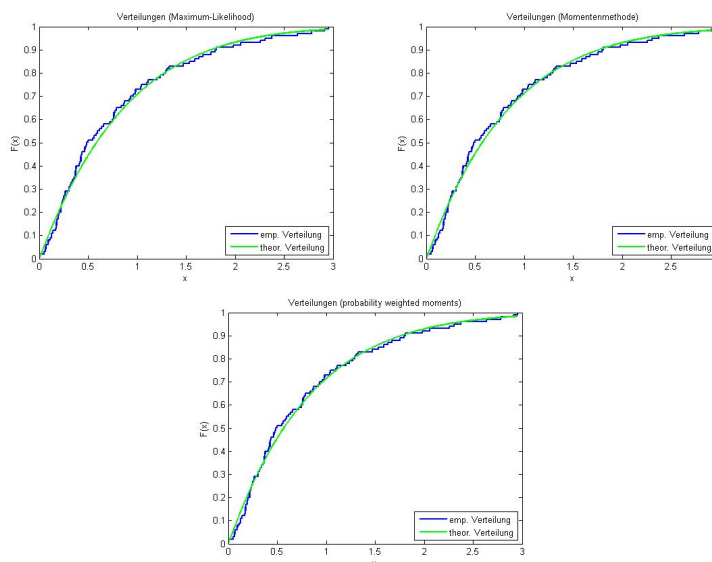


Abbildung 4.10: Vergleich der Verteilungen mit Maximum-Likelihood, Momentenmethode und wahrscheinlichkeitsgewichteten Momenten

Hier ist zu erkennen, dass die Abweichung zwischen empirischer und geschätzter Verteilung bei Maximum-Likelihood bei einem Wert von ca. 0,5 und 2,25 am größten ist. Trotz allem ergeben sich ähnliche Kenngrößen:

Schätzverfahren	Kolmogorov-Smirnov-Test	p -Level	$VaR_{99\%}$
Maximum-Likelihood	0	0,7614	5,55
Momentenmethode	0	0,8142	5,55
PWM	0	0,7803	5,55

Zusammenfassend kann man sagen, dass die Erhöhung der Schwelle einen Anstieg des p -Levels hervorgerufen hat, was laut Theorie zu erwarten war. Allerdings ergaben sich bei höherer Schranke u auch mehrere Abweichungen bei den Quantilen. Sonst waren jedoch keine signifikanten Unterschiede speziell zwischen Maximum-Likelihood und wahrscheinlichkeitsgewichteten Momenten zu erkennen.

Autoregressive Zeitreihen mit Sprüngen

Wir modifizieren nun den vorherigen autoregressiven Prozess und fügen eine Sprungkomponente in die Bildungsvorschrift hinzu, um plötzliche Preissprünge zu simulieren. Als Schwellenwert u wird das 0,91-Quantil der Daten benutzt:

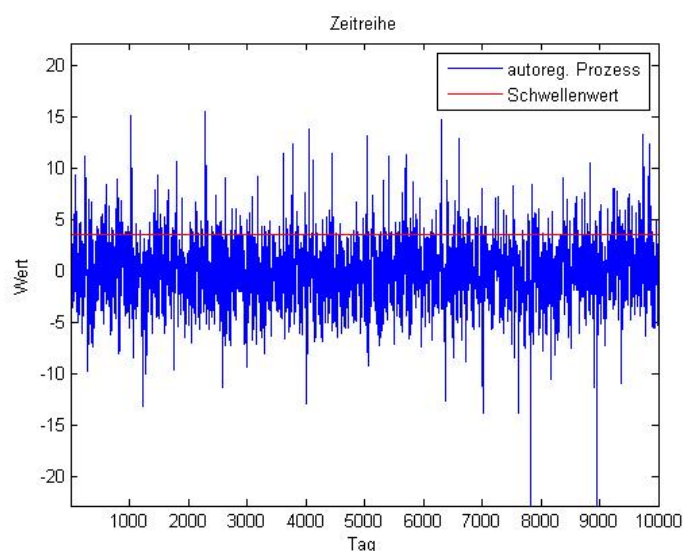


Abbildung 4.11: autoregressiver Prozess mit Jumps über 10.000 Tage mit 0,91-Quantil der Daten als Schwellenwert

Die Schwelle u liegt hier bei dem 0,91-Quantil der Daten. Für die Parameterschätzung werden die drei verschiedenen Schätzmethoden angewendet. Es ergeben sich folgende QQ-Plots:

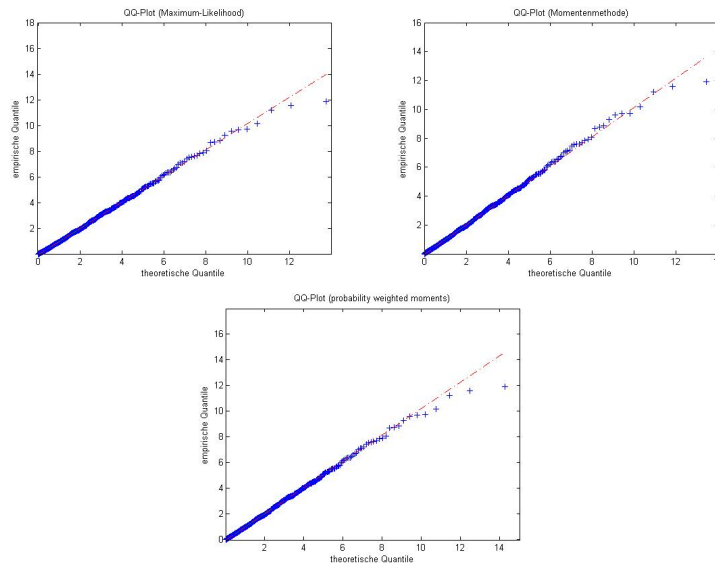


Abbildung 4.12: Vergleich der QQ-Plots mit Maximum-Likelihood, Momentenmethode und wahrscheinlichkeitsgewichteten Momenten

In der Grafik ist zu erkennen, dass alle drei QQ-Plots in etwa die gleichen Quantile hervorbringen. Das heißt, dass die drei verschiedenen Parameterschätzungen näherungsweise die gleichen Parameter berechnen. Bei den drei höchsten Quantilen sind minimale Unterschiede zu erkennen, die jedoch nicht signifikant sind.

Bei der Betrachtung der Verteilungsfunktion ergeben sich folgende Bilder:

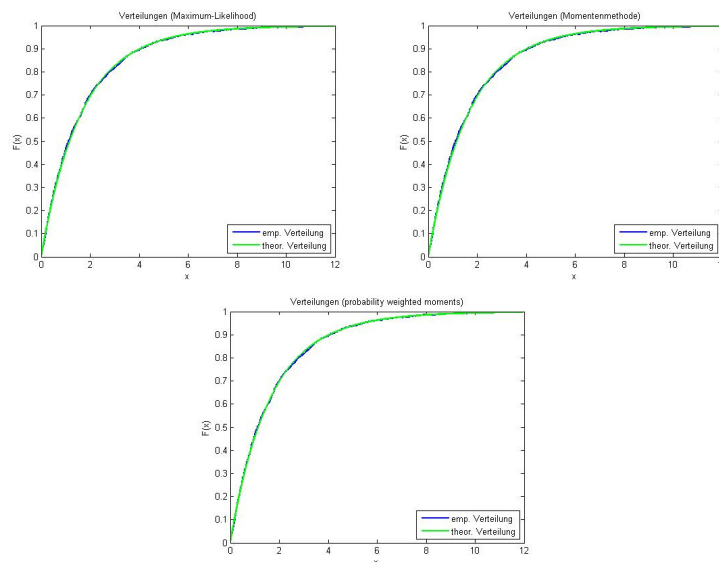


Abbildung 4.13: Vergleich der Verteilungen mit Maximum-Likelihood, Momentenmethode und wahrscheinlichkeitsgewichteten Momenten

Auch hier erkennt man deutlich, dass die verschiedenen Parameterschätzungen keine signifikanten Unterschiede bei der Verteilung hervorrufen. Dies lässt sich auch in der folgenden Tabelle der Kenngrößen beobachten:

Schätzverfahren	Kolmogorov-Smirnov-Test	p -Level	$VaR_{99\%}$
Maximum-Likelihood	0	0,9803	7,45
Momentenmethode	0	0,9464	7,45
PWM	0	0,9992	7,45

Als nächstes wollen wir nun die Schranke u erhöhen, um zu sehen, wie sich Quantile und Verteilung mit weniger und extremeren Daten verhalten:

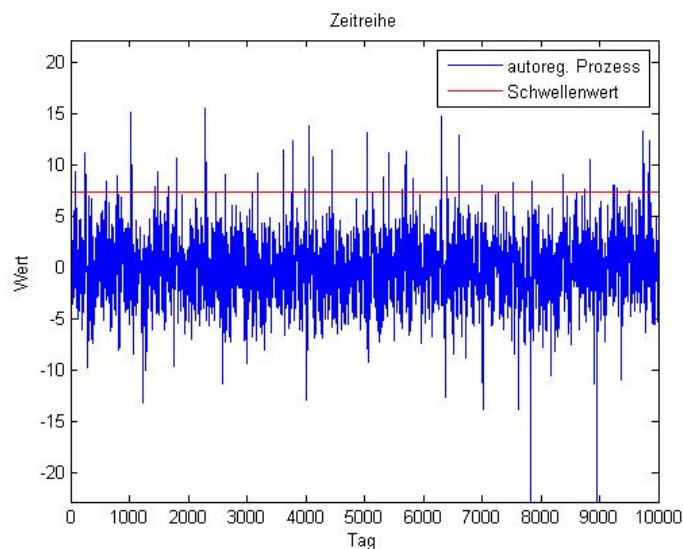


Abbildung 4.14: autoregressiver Prozess mit Jumps über 10.000 Tage mit 0,99-Quantil der Daten als Schwellenwert

Der Schwellenwert hat sich nun deutlich erhöht auf ca. 7. Wir wenden wieder die drei verschiedenen Parameterschätzer an. Es ergeben sich folgende QQ-Plots:

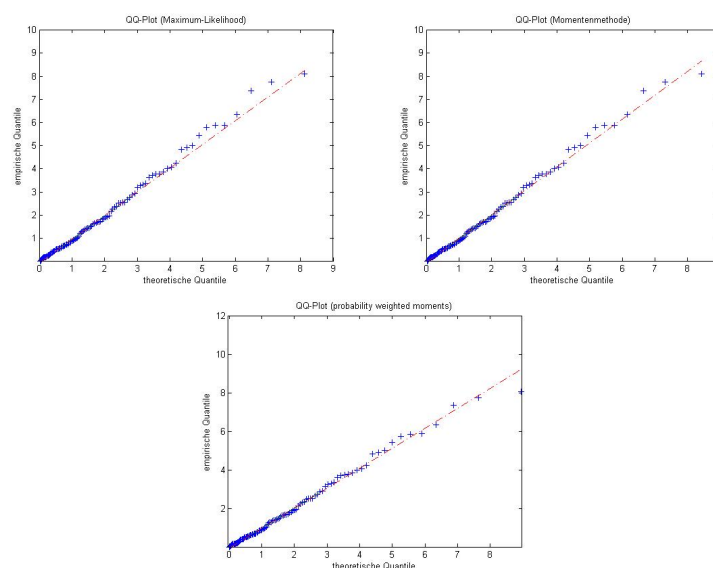


Abbildung 4.15: Vergleich der QQ-Plots mit Maximum-Likelihood, Momentenmethode und wahrscheinlichkeitsgewichteten Momenten

Im Vergleich zu den QQ-Plots mit kleinerer Schranke u , liegen hier die drei höchsten Quantile näher an der Winkelhalbierenden, sodass Sie den empirischen Quantilen näher sind als zuvor. Ansonsten sind die QQ-Plots mit den verschiedenen Schätzmethoden nahezu identisch.

Untersuchen wir nun die Verteilungsfunktion:

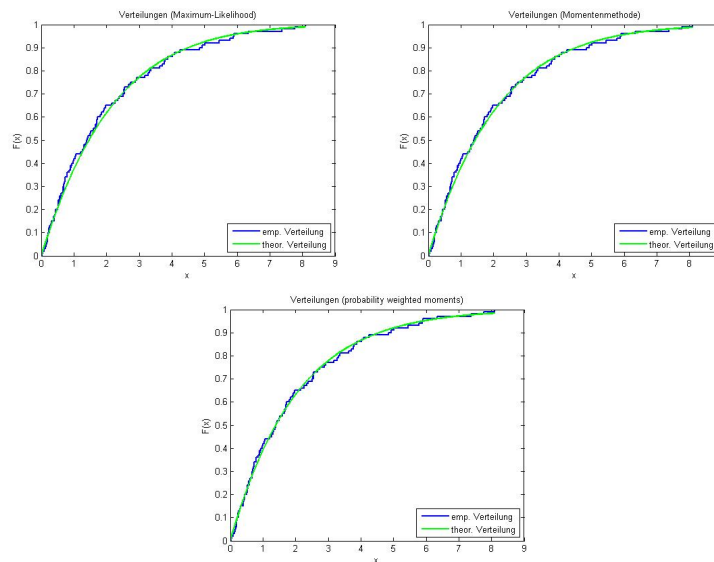


Abbildung 4.16: Vergleich der Verteilungen mit Maximum-Likelihood, Momentenmethode und wahrscheinlichkeitsgewichteten Momenten

Das Verhalten der Verteilungsfunktion ist bei allen drei Schätzmethoden ähnlich. Im unteren Bereich der x -Achse sind leichte Unterschiede zu erkennen, bei denen die Verteilungsfunktion mit der Maximum-Likelihood-Schätzung minimal von der empirischen Verteilung abweicht. Darüber hinaus sind jedoch keine signifikanten Unterschiede zu erkennen.

Schätzverfahren	Kolmogorov-Smirnov-Test	p -Level	$VaR_{99\%}$
Maximum-Likelihood	0	0,9363	7,47
Momentenmethode	0	0,9788	7,47
PWM	0	0,9986	7,47

In der obigen Tabelle ist zu erkennen, dass die Güte der Verteilungsanpassung mit allen drei Schätzverfahren hervorragend ist. Mit jeder Schätzmethode erreichen wir ein p -Level von über 90 Prozent und errechnen den gleichen Value-at-Risk.

Zusammenfassend ist zu sagen, dass alle drei Methoden der Parameterschätzung hier in etwa dieselbe Güte haben. Die Erhöhung der Schwelle u brachte jedoch keinen Gewinn im p -Level, wobei die Quantile im höheren Bereich leicht besser geschätzt wurden.

4.5.2 Preiszeitreihen

Phelix-Index Baseload

Nun wollen wir die Theorie der Schwellenwertüberschreitung an echten Daten aus der Praxis anwenden. Zunächst wird der Phelix-Index im Baseload betrachtet und die Schwelle u auf das 0,91-Quantil der Daten gesetzt:

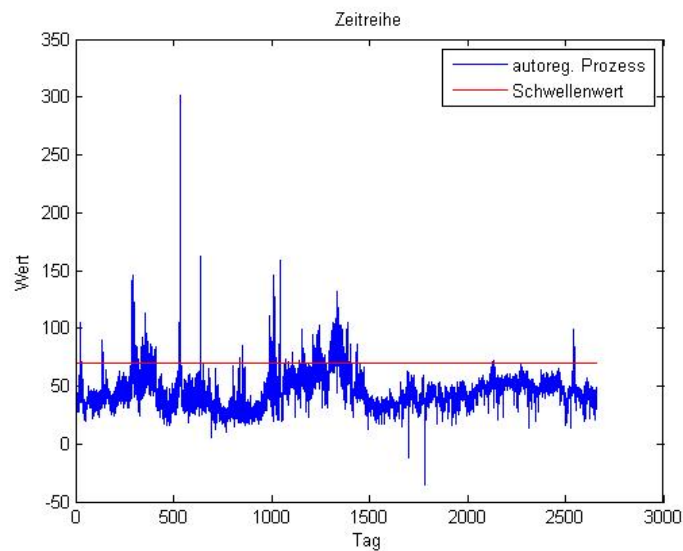


Abbildung 4.17: Zeitreihe des Phelix-Index im Baseload über 2662 Tage mit 0,91-Quantil der Daten als Schwellenwert

Wie auch bei unseren simulierten Zeitreihen, wollen wir untersuchen, wie sich die drei Parameterschätzer verhalten und welche Ergebnisse diese liefern. Beginnen wir mit dem Vergleich der Quantile:

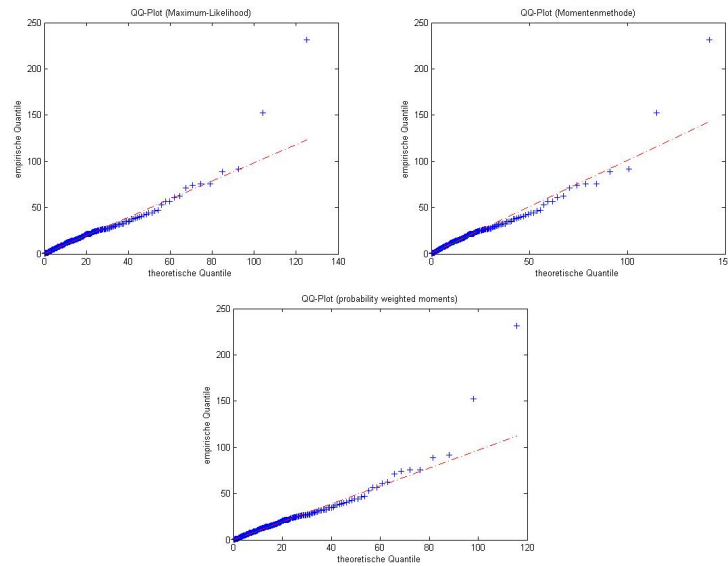


Abbildung 4.18: Vergleich der QQ-Plots mit Maximum-Likelihood, Momentenmethode und wahrscheinlichkeitsgewichteten Momenten

Im näheren Vergleich stellen wir fest, dass die Quantile mit der Maximum-Likelihood-Schätzung am besten getroffen werden. Bei den anderen zwei Schätzmethode kommt es im Bereich der höheren Quantile zu einer leichten Unter- bzw. Überschätzung. Die zwei größten Quantile sind jedoch mit allen drei Verfahren nicht erreichbar.

Bei der Verteilung sieht man jedoch bei allen drei Verfahren den gleichen Verlauf. Es gibt kleine Abweichungen im Wertebereich zwischen den x -Werten 30 und 60:

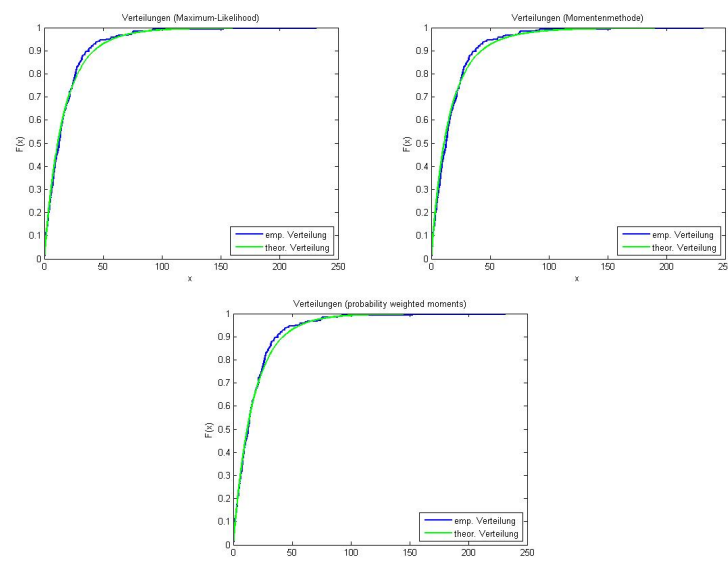


Abbildung 4.19: Vergleich der Verteilungen mit Maximum-Likelihood, Momentenmethode und wahrscheinlichkeitsgewichteten Momenten

Diese Abweichung macht sich beim p -Level bemerkbar:

Schätzverfahren	Kolmogorov-Smirnov-Test	p -Level	$VaR_{99\%}$
Maximum-Likelihood	0	0,5669	110,28 EUR/MWh
Momentenmethode	0	0,2505	110,44 EUR/MWh
PWM	0	0,7533	110,46 EUR/MWh

Der $VaR_{99\%}$ bleibt bei allen Schätzverfahren auf dem gleichen Niveau. Nur das p -Level schwankt zwischen 25 und 75 Prozent, was auf den oben genannten Unterschied im Bereich zwischen 30 und 60 EUR/MWh zurückzuführen ist.

Wir erhöhen nun wieder den Schwellenwert u . Da bei den Preiszeitreihen deutlich weniger Daten zur Verfügung stehen, hätten wir bei dem 0,99-Quantil nur 26 Datenpunkte zur Verfügung, dessen Überschreitungen größer als 0 ist. Dieser Datenumfang ist jedoch für eine Verteilungsannahme ungeeignet. Deshalb wird das 0,98-Quantil der Daten benutzt, sodass uns noch 53 Datenpunkte erhalten bleiben:

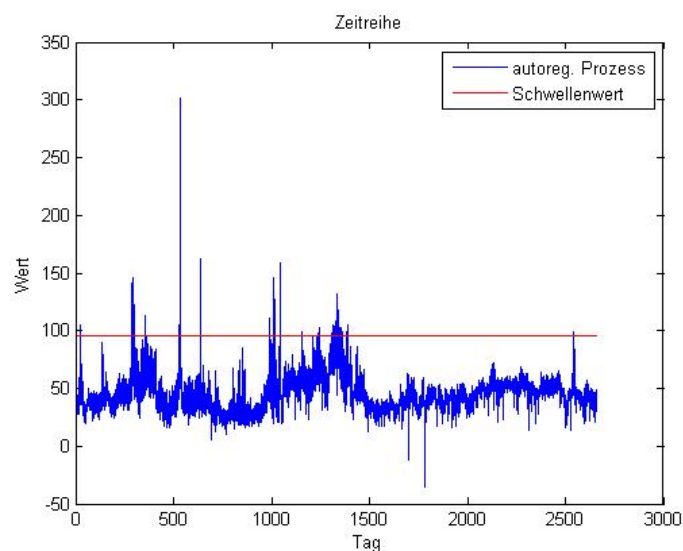


Abbildung 4.20: Zeitreihe des Phelix-Index im Baseload über 2662 Tage mit 0,98-Quantil der Daten als Schwellenwert

Mit den drei Schätzverfahren ergeben sich folgende Quantile:

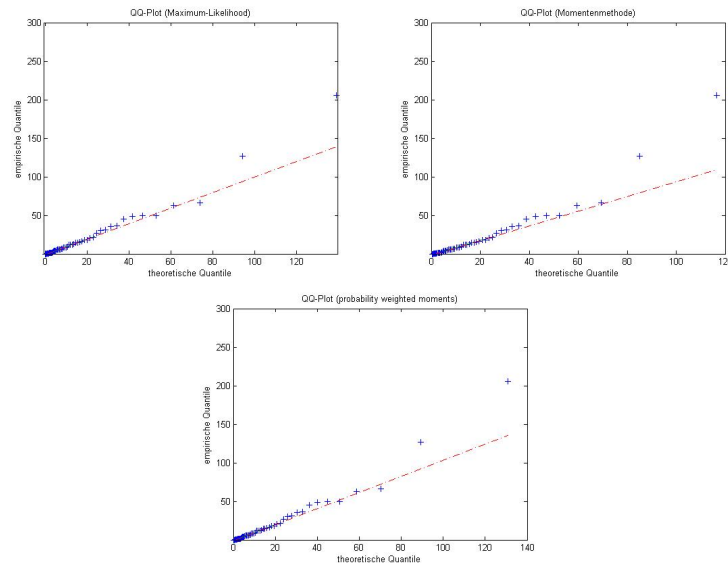


Abbildung 4.21: Vergleich der QQ-Plots mit Maximum-Likelihood, Momentenmethode und wahrscheinlichkeitsgewichteten Momenten

Die QQ-Plots Maximum-Likelihood und Momentenmethode sind nahezu identisch. Die Quantile mit der Momentenmethode liegen minimal über der Winkelhalbierenden. Darüber hinaus weisen die Quantile aller Schätzmethoden die gleiche Struktur auf. Erst bei einer Überschreitung von 150 EUR/MWh treten deutliche Abweichungen auf, die auch schon mit dem Schwellenwert des 0,91-Quantils aufgetreten sind.

Bei der Verteilung kann jedoch zwischen den verschiedenen Schätzungen kein signifikanter Unterschied festgestellt werden:

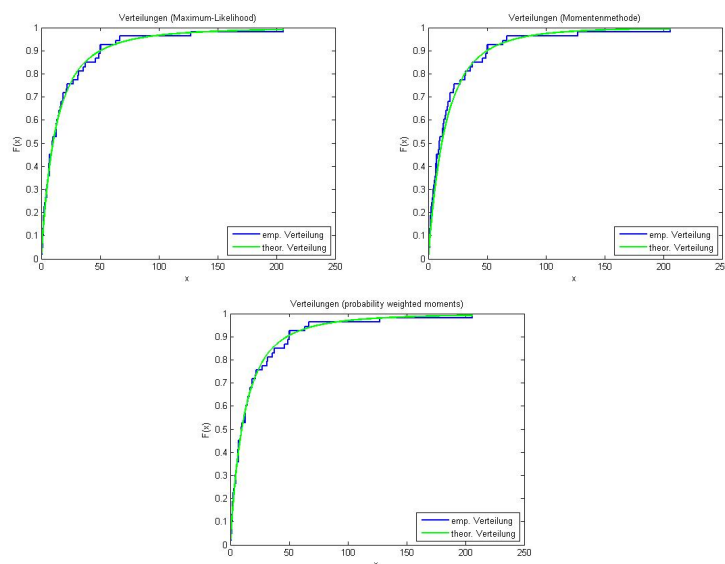


Abbildung 4.22: Vergleich der Verteilungen mit Maximum-Likelihood, Momentenmethode und wahrscheinlichkeitsgewichteten Momenten

Es ergeben sich folgende Kenngrößen:

Schätzverfahren	Kolmogorov-Smirnov-Test	p -Level	$VaR_{99\%}$
Maximum-Likelihood	0	0,9926	105,45 EUR/MWh
Momentenmethode	0	0,7945	106,93 EUR/MWh
PWM	0	0,9964	105,20 EUR/MWh

Im Vergleich zur vorherigen niedrigeren Schranke u hat sich das p -Level deutlich verbessert. Auffällig ist, dass der VaR_{99} hier deutlich kleiner ist als zuvor. Da aber zwischen den Preisen von 110 und 105 EUR/MWh fast keine weiteren Preise gemessen wurden, liegt die Wahrscheinlichkeit einer Überschreitung beider Werte bei maximal 1 Prozent.

Gesamt gesehen ergaben die drei verschiedenen Schätzmethoden die gleichen Ergebnisse, obwohl mit der Momentenmethode ein deutlich schlechteres p -Level einherging. Die Erhöhung der Schwelle u brachte bei allen Schätzverfahren einen Anstieg des p -Levels.

Phelix-Index Peakload

Betrachten wir zum Schluss noch den Peakload des Phelix-Index, zunächst noch mit dem 0,91-Quantil der Daten als Schwellenwert:

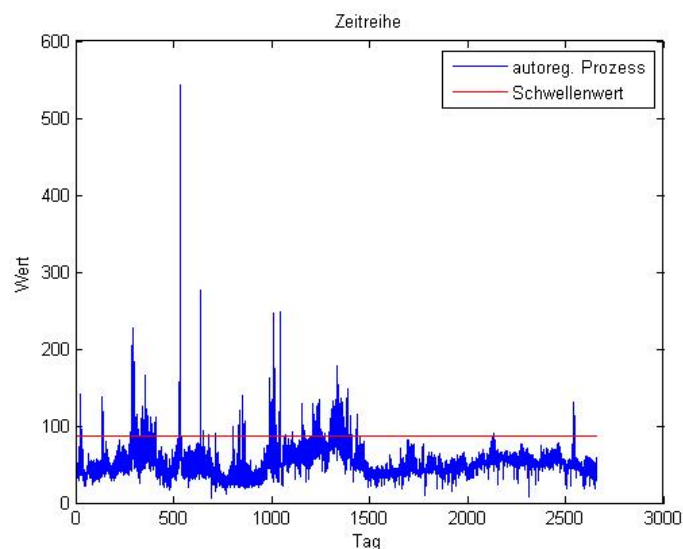


Abbildung 4.23: Zeitreihe des Phelix-Index im Peakload über 2662 Tage mit 0,91-Quantil der Daten als Schwellenwert

Wir wollen wie immer untersuchen, wie sich die drei Parameterschätzer verhalten und welche Ergebnisse diese liefern. Beginnen wir mit dem Vergleich der Quantile:

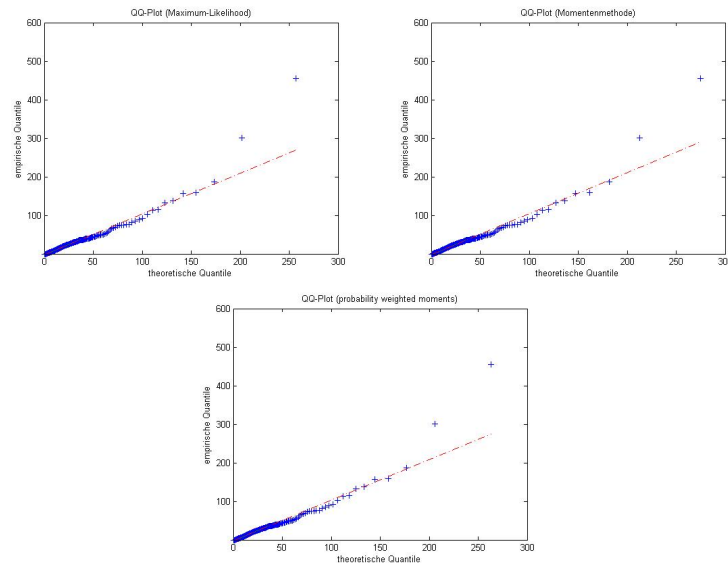


Abbildung 4.24: Vergleich der QQ-Plots mit Maximum-Likelihood, Momentenmethode und wahrscheinlichkeitsgewichteten Momenten

Im näheren Vergleich stellen wir fest, dass die Quantile mit allen Schätzverfahren gleich getroffen werden. Die zwei größten Quantile sind jedoch, wie schon beim Baseload mit allen drei Verfahren nicht erreichbar.

Bei der Verteilung sieht man ebenfalls bei allen drei Verfahren den gleichen Verlauf:

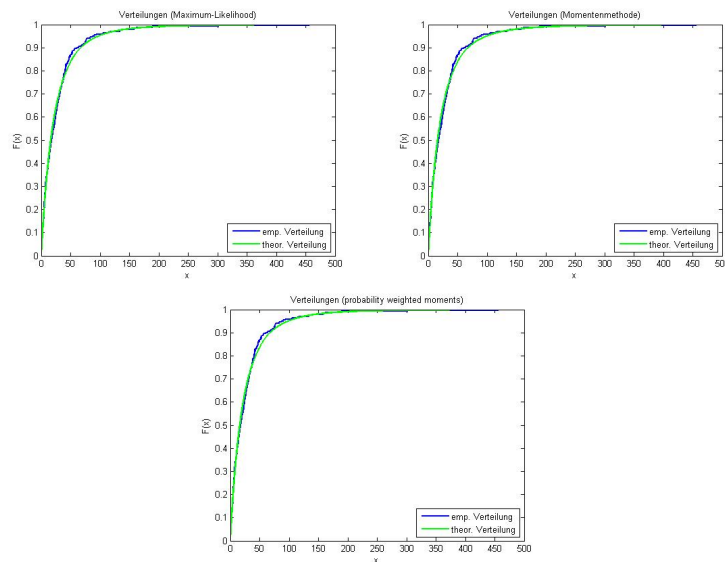


Abbildung 4.25: Vergleich der Verteilungen mit Maximum-Likelihood, Momentenmethode und wahrscheinlichkeitsgewichteten Momenten

Als Zusammenfassung sehen wir hier die Kenngrößen:

Schätzverfahren	Kolmogorov-Smirnov-Test	p -Level	VaR_{99}
Maximum-Likelihood	0	0,7590	150,63 EUR/MWh
Momentenmethode	0	0,6309	150,97 EUR/MWh
PWM	0	0,7870	151,20 EUR/MWh

Der VaR_{99} bleibt bei allen Schätzverfahren auf dem gleichen Niveau. Nur das p -Level schwankt zwischen 60 und 80 Prozent.

Wir erhöhen nun wieder den Schwellenwert u . Benutzt wird abermals das 0,98-Quantil, um einen gewissen Datenumfang zu bewahren:

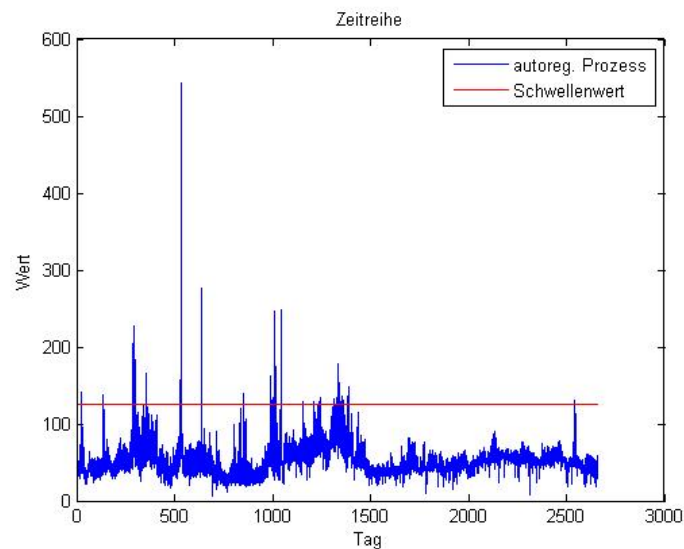


Abbildung 4.26: Zeitreihe des Phelix-Index im Peakload über 2662 Tage mit 0,98-Quantil der Daten als Schwellenwert

Mit den drei Schätzverfahren ergeben sich folgende Quantile:

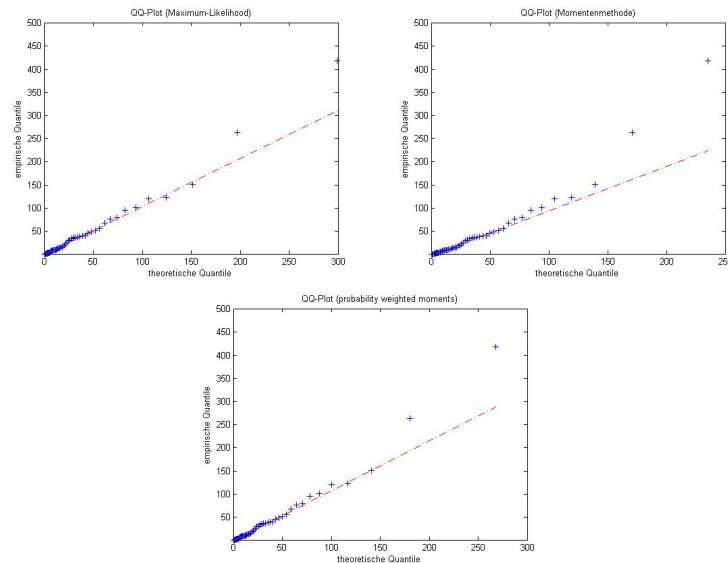


Abbildung 4.27: Vergleich der QQ-Plots mit Maximum-Likelihood, Momentenmethode und wahrscheinlichkeitsgewichteten Momenten

Die QQ-Plots der Maximum-Likelihood-Methode und der wahrscheinlichkeitsgewichteten Momente sind nahezu identisch. Die Quantile mit der Momentenmethode liegen minimal über der Winkelhalbierenden. Darüber hinaus weisen die Quantile aller Schätzmethoden die gleiche Struktur auf. Erst bei einer Überschreitung von 250 EUR/MWh treten deutliche Abweichungen auf, die auch schon mit dem Schwellenwert des 0,91-Quantils erkennbar waren.

Auch bei den Verteilungen kann zwischen Maximum-Likelihood und wahrscheinlichkeitsgewichteten Momenten kein signifikanter Unterschied festgestellt werden. Nur die Verteilung mit Momentenmethode weist kleine Abweichungen auf:

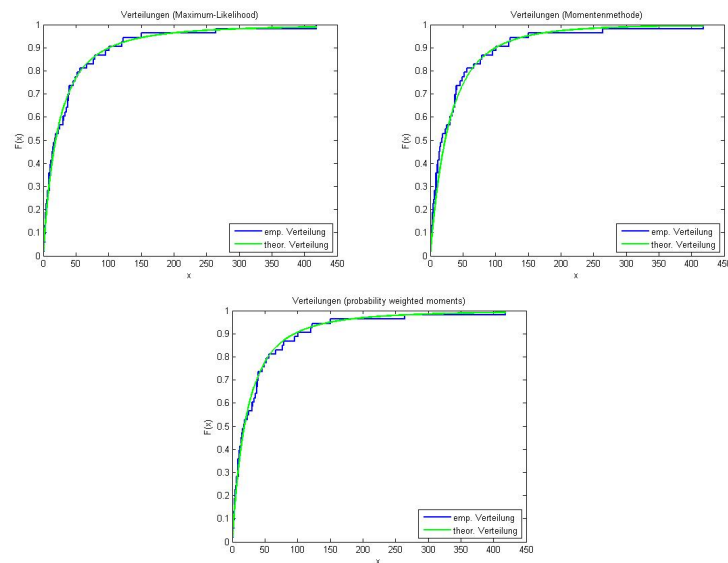


Abbildung 4.28: Vergleich der Verteilungen mit Maximum-Likelihood, Momentenmethode und wahrscheinlichkeitsgewichteten Momenten

Es ergeben sich folgende Kenngrößen:

Schätzverfahren	Kolmogorov-Smirnov-Test	p -Level	VaR_{99}
Maximum-Likelihood	0	0,9486	144,46 EUR/MWh
Momentenmethode	0	0,5555	148,26 EUR/MWh
PWM	0	0,8985	144,15 EUR/MWh

Im Vergleich zur vorherigen niedrigeren Schranke u hat sich das p -Level, außer bei der Momentenmethode, deutlich verbessert. Auffällig ist auch hier wieder, dass der VaR_{99} etwas kleiner ist als zuvor. Da aber zwischen den Preisen von 150 und 144 EUR/MWh fast keine weiteren Preise gemessen wurden, liegt die Wahrscheinlichkeit einer Überschreitung beider Werte bei maximal 1 Prozent.

Zusammenfassend ist zu sagen, dass der Maximum-Likelihood-Schätzer und die wahrscheinlichkeitsgewichteten Momente nahezu identische Ergebnisse lieferten. Die Erhöhung der Schwelle u ergab einen deutlichen Anstieg des p -Levels bei beiden Verfahren. Nur mit der Momentenmethode wurden schlechtere p -Levels erreicht und es ergaben sich auch größere Abweichungen bei den Quantilen, als bei den anderen zwei Verfahren. Somit ergeben sich die wahrscheinlichkeitsgewichteten Momente als wirkliche Alternative zur Maximum-Likelihood-Schätzung, da bei Letzterem im ungünstigen Fall ein nichtlineares Gleichungssystem gelöst werden muss.

5 Zusammenfassung und Ausblick

Das Ziel dieser Arbeit bestand darin, extreme Strompreise am Spotmarkt der EEX Leipzig zu analysieren und zu untersuchen, ob diesen seltenen Ereignissen eine Wahrscheinlichkeitsverteilung zugrunde liegt. Mit zwei bekannten Verfahren der Extremwerttheorie wurde eine ausführliche Untersuchung durchgeführt und die entstandenen Ergebnisse diskutiert. Dabei wurde auf theoretische, als auch auf praktische Aspekte eingegangen. Das Ergebnis dieser Untersuchung ist, dass für die Anwendung der Verfahren mehrere Probleme auftauchen. Zunächst sind die Zeitreihen vom Spotmarkt nur näherungsweise stationär. Wie in Kapitel 3.4 dargestellt wurde, ist die Stationarität nur eine hinreichende und keine notwendige Bedingung. Dennoch kann im Allgemeinen bei Nichtstationarität ein Qualitätsverlust der Ergebnisse angenommen werden. Betrachten wir die Resultate aus Kapitel 3.7 und 4.5, so kann bei beiden Verfahren festgestellt werden, dass die kleineren Quantile der geschätzten Verteilung zwar mit den Daten übereinstimmen, aber die großen Quantile, die uns eigentlich interessieren, nicht getroffen wurden. Dies hängt mit dem noch zu kleinen Datenumfang der Strombörse zusammen. Während den aufgezeichneten 2662 Börsentagen sind einige Ereignisse dabei, die als Extremwerte identifiziert werden können. Die Anzahl dieser extrem hohen Preise ist jedoch noch stark begrenzt. Zudem werden Unterschiede zwischen den Ergebnissen beider Verfahren deutlich. Die Quantile bei Peak over Threshold stimmen häufiger mit denen der Daten überein, was darauf zurückzuführen ist, dass bei der Theorie der Schwellenwertüberschreitung alle Daten, die eine gewisse Schranke überschreiten, mit in die Schätzung einfließen. Peak over Threshold ist somit bei geringem Datenumfang zu bevorzugen. Die Theorie der Blockmaxima hat jedoch den Vorteil, dass dort wirkliche Strompreise betrachtet werden, während bei Peak over Threshold nur Exzesse eine Rolle spielen. Nachteil der Blockmaximatheorie ist jedoch die große Datenverschwendung, da nur ein einziger Wert aus jedem Block gezogen wird. Treten innerhalb eines Blocks mehrere extreme Strompreise auf, so wird dennoch nur der größte herausgenommen und die anderen verworfen. Dies stellt jedoch keinen effizienten Umgang mit Informationen dar. Des Weiteren ist die Wahl der Blocklänge ein Problem. Wie wir in Kapitel 3.6 gesehen haben, ist es sehr schwierig, eine allgemeine Aussage zur optimalen Blocklänge einer Zeitreihe zu treffen. Die Länge eines Blocks ist aber im Hinblick auf den Datenumfang sehr entscheidend. Kleine Blöcke implizieren viele nutzbare Daten. Allerdings treten in kurzen Zeitperioden nicht immer Extremwerte auf. Wählt man die Blöcke zu groß, so schrumpft die Menge der Daten, an die unsere Extremwertverteilung angepasst werden soll erheblich, so dass man zu keinem statistisch glaubwürdigen Ergebnis kommen kann.

Ziel war es außerdem, den Value at Risk mit beiden Verfahren abzuschätzen. Hierbei kamen wir ebenfalls zu unterschiedlichen Ergebnissen. Grund dafür war, dass der Value at Risk der Blockmaxima den Maximalpreis angibt, der innerhalb eines Blocks mit

einer gegebenen Wahrscheinlichkeit nicht überschritten wird. Die Theorie der Schwellenwertüberschreitung berechnet dagegen das p -Quantil des Tagespreises. Der Value at Risk der Blockmaxima ist deshalb in der Hinsicht weitaus sicherer, da es sich um das p -Quantil der Verteilung der Maxima handelt. Andererseits fordert diese Berechnung eine große Menge von Daten, die zur Zeit an der Leipziger Strombörse, auf Grund der kurzen Zeit, die nach der Gründung verstrichen ist, nicht vorhanden ist.

Das momentane Problem des geringen Datenumfangs wird sich allerdings in den nächsten Jahren langsam verbessern, da im Laufe der Zeit die aufgezeichneten Zeitreihen immer länger werden. Dies führt zu mehr Informationen über zukünftige Extremwerte und somit zu besseren Ergebnissen bei der Anwendung der Extremwerttheorie.

Weitere interessante Aspekte sind unter anderem alternative Risikomaße, wie das Return-Level [5], bei dem die Häufigkeit von hohen Preisen beobachtet und dann deren Ausmaß geschätzt wird oder die Return-Period [5], die die Höhe der Preise beobachtet und dann deren Häufigkeit schätzt. Des Weiteren kann man die gleichen Untersuchungen auch auf dem Terminmarkt tätigen. Dazu müssen beide Verfahren getrennt voneinander angepasst werden, da es sich dann nicht mehr zwangsläufig um stationäre Zeitreihen handelt. Grund dafür sind mögliche globale Trends in der Zeitreihe. Darüber hinaus treten am Terminmarkt Extremwerte häufiger in Clustern auf, was wir in dieser Arbeit beim Spotmarkt ausgeschlossen haben.

Anhang A: Kwiatkowski-Phillips-Schmidt-Shin-Test

Der Kwiatkowski-Phillips-Schmidt-Shin-Test, kurz KPSS-Test, ist ein Hypothesen-Test aus der schließenden Statistik, der direkt auf (Trend-)Stationarität testet. Zunächst benötigen wir folgende Definition [11]:

Definition A.1 Ein stochastischer Prozess $(X_t)_{t \in L}$, mit L als Indexmenge, heißt **integriert der Ordnung d** , falls $(X_t)_{t \in L}$ d -mal differenziert werden muss, damit der Prozess stationär wird.

Die Idee des KPSS-Tests liegt darin, dass sich jeder integrierte Prozess in der folgenden Form darstellen lässt:

$$X_t = \alpha + \delta t + d \sum_{i=1}^t Z_i + U_t,$$

mit U_t als stationären Prozess, $\alpha + \delta t$ als lineare Trendkomponente und $\sum_{i=1}^t Z_i$ als Random-Walk. Man kann in diesem Test wählen, ob mit oder ohne linearem Trend getestet werden soll. Bei unseren Zeitreihen gehen wir davon aus, dass kein linearer Trend vorliegt. Ist $d = 0$, so handelt es sich um einen trendstationären Prozess. Somit ergeben sich folgende Hypothesen:

$$H_0 : d = 0$$

$$H_1 : d \neq 0.$$

Als erstes wird eine Regression von X_t auf eine Konstante und, falls notwendig, auf einen Zeittrend mittels der Methode der kleinsten Quadrate unter der Nullhypothese $d = 0$ durchgeführt. Damit erhalten wir die Residuen \hat{u}_t , bzw die Abweichung zwischen der Regressionsgeraden und dem Prozess $(X_t)_{t \in L}$, sowie deren Partialsumme $S_T = \sum_{t=1}^T \hat{u}_t$, mit T als Zeitreihenlänge. Die Teststatistik lautet somit [11]:

$$KPSS_T = \frac{\sum_{t=1}^T S_t^2}{T^2 \hat{\omega}_T^2}.$$

Dabei ist $\hat{\omega}_T^2$ der Schätzer für die langfristige Varianz:

$$\begin{aligned}\hat{\omega}_T^2 &= \hat{\sigma}_u^2 + 2 \sum_{\tau=1}^A \left(1 - \frac{\tau}{T-1}\right) \hat{\gamma}_\tau \\ \hat{\gamma}_\tau &= \frac{1}{T} \sum_{t=1}^{T-\tau} (\hat{u}_t - \bar{u})(\hat{u}_{t+\tau} - \bar{u}) \\ \hat{\sigma}_u^2 &= \frac{1}{T-1} \sum_{t=1}^T (\hat{u}_t - \bar{u})^2,\end{aligned}$$

mit $A \leq T$ als Abbruchwert der Lags. Anschließend wird der Wert der Teststatistik mit dem $(1 - \alpha)$ -Quantil zum Konfidenzniveau α der KPSS-Verteilung verglichen.

Anhang B: Kolmogorov-Smirnov-Test

Der Kolmogorov-Smirnov-Test ist ein Hypothesen-Test auf Verteilung und untersucht eine vorher festgelegte Verteilungsannahme. Die Nullhypothese geht davon aus, dass die vorliegenden n Daten Realisierungen eines Vektors von Zufallsgrößen mit Verteilungsfunktion F sind. Die Teststatistik sieht wie folgt aus [7]:

$$T = \sqrt{n} \sup_x |F_n(x) - F(x)|,$$

mit $F_n(x)$ als empirische Verteilungsfunktion der vorliegenden Daten. Unter H_0 besitzt T für $n \rightarrow \infty$ eine Kolmogorov-Verteilung. Die Nullhypothese wird zum Niveau $1 - \alpha$ verworfen (im Allgemeinen wird $\alpha = 0,05$ gesetzt), wenn $T > K_{1-\alpha}$, wobei $K_{1-\alpha}$ das $1 - \alpha$ -Quantil der Kolmogorov-Verteilung ist. Sonst ist gegen H_0 nichts einzuwenden. Als Gütemaß für den Kolmogorov-Smirnov-Test gibt es das p -Level $p \in [0, 1]$, welches uns sagt, wie wahrscheinlich es ist, dass die zu Grunde liegenden Daten wirklich zu der angepassten Verteilung gehören [16].

Literaturverzeichnis

- [1] McNeil, A.J.; Frey, R.; Embrechts, P.: Quantitative Risk Management: Concepts Techniques Tools, Princeton University Press, 2005
- [2] Prof.Dr. Dette, H.; Dr. Birke, Melanie: Einführung in die Wahrscheinlichkeitstheorie und Statistik, Ruhr-Universität Bochum, Fakultät für Mathematik, Lehrstuhl für Stochastik, <http://www.ruhr-uni-bochum.de/imperia/md/content/mathematik3/lehre/ws0809/blatt7.pdf>
- [3] Löwe, Matthias: Extremwerttheorie, Universität Münster, Institut für Statistik, Skript, <http://wwwmath.uni-muenster.de/statistik/loewe/extrem.pdf>
- [4] Woeste, Birgit: Eine Anwendung der Blockmaxima Methode im Risikomanagement, Westfälische Wilhelms-Universität Münster, Institut für Statistik, Diplomarbeit
- [5] Embrechts, P.; Klüppelberg C.; Mikosch T.: Modeling Extremal Events, for Insurance and Finance, 1. Auflage 1996, Springer Verlag
- [6] Prahm, Julia: Eine Anwendung der Peak over Threshold Methode im Risikomanagement, Westfälische Wilhelms-Universität Münster, Institut für Statistik, Diplomarbeit
- [7] Hölzswimmer, Manuela: Extremwerte II - Peak over Threshold (POT), LMU München, Interdisziplinäres Seminar, Statistische Verfahren in den Geowissenschaften, Belegarbeit, 2009
- [8] Fischer, Matthias: Extremwertstatistik mit Anwendungen in Finanzmärkten, Versicherungsmärkten und der Klimatologie, Universität Erlangen-Nürnberg, Vorlesungsskript, 2010
- [9] Frahm, Gabriel: Ermittlung des Value-at-Risk von Finanzportefeuilles mit Methoden der Extremwerttheorie, Universität Köln, Diplomarbeit
- [10] Schlittgen, R.: Angewandte Zeitreihenanalyse, Oldenbourg Wissenschaftsverlag GmbH, 2001, ISBN: 3-486-25805-2
- [11] Winkler, Angelika: Anwendung nichtlinearer Volatilitätsmodelle zur Bemessung von Sicherheitsanforderungen bei der European Commodity Clearing AG, Hochschule für Technik, Wirtschaft und Kultur Leipzig, Bachelorarbeit
- [12] Kern, Sigrid Gerlinde: Die Stochastische Modellierung des EEX-Spotmarktes und die Bewertung von Swing-Optionen, Technische Universität Graz, Institut für Statistik, Dissertation
- [13] Tankow, Peter; Voltchkova, Ekaterina: Jump-diffusion models: a practitioner's guide, Universität Paris, Universität Toulouse, wissenschaftlicher Artikel

- [14] Kabluchko, Zakhar: Extremwerttheorie, Universität Ulm, Fakultät für Mathematik und Wirtschaftswissenschaften, Vorlesungsskript, 2011
- [15] Bauer, Heinz: Maß- und Integrationstheorie, 2. Auflage 1992, De Gruyter Verlag
- [16] Hatzinger, R.; Hornik, K.; Nagel, H.: R: Einführung durch angewandte Statistik, 1. Auflage 2011, Pearson Studium Verlag
- [17] Hosking, J.R.M.; Wallis, J.R.: Parameter and Quantile Estimation for the Generalized Pareto Distribution, Technometrics, August 1987, Vol. 29, No. 3, wissenschaftlicher Artikel

Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Stellen, die wörtlich oder sinngemäß aus Quellen entnommen wurden, sind als solche kenntlich gemacht.

Diese Arbeit wurde in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegt.

Mittweida, 23. September 2012